NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

QUALITY OF USMC OFFICERS: BUILDUP VS. REDUCTION IN FORCES

by

Michael S. Griner

March 2016

Thesis Advisor: Marigee Bacolod
Second Reader: Chad Seagren

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704–0188*

1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** March 2016	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE QUALITY OF USMC OFFICERS: BUILDUP VS. REDUCTION IN FORCES	**5. FUNDING NUMBERS**
6. AUTHOR(S) Michael S. Griner	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ____NPS.2015.0044-IR-EP5-A___.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited	**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

At the epicenter of maintaining the finest, most professional, and most feared military in the world is the quality of its leaders. The United States Marine Corps' Officer Corps is the standard bearer, with an accession pipeline more rigorous than any in the Department of Defense. As is the historical norm, however, the U.S. military undertakes ambitious increases to end strength in the wake of a prolonged conflict. As a conflict fades, the increased end strength is no longer warranted. The typical approach during the drawdown is to reduce accessions, create stricter retention policies, and entice members to leave the service through voluntary measures.

This research identifies the trade-off between quantity and quality necessitated by end-strength changes. Quantitative analysis using a difference-in-differences research design shows, relative to the buildup, officer quality increases during the drawdown. In particular, combat and non-combat occupations have a 0.0321 and 0.0834 point increase, respectively, in FITREP scores in the drawdown compared to the control group. Alternative measures and additional robustness checks support the hypothesis that the drawdown yields higher-quality officers. It is imperative to adapt force-shaping policies to a gradual approach to ensure the Marines retained during the buildup meet the quality standards of the drawdown.

14. SUBJECT TERMS United States Marine Corps, fitness report, end strength, buildup, reduction in forces, performance evaluation system	**15. NUMBER OF PAGES** 89
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540–01-280-5500

Standard Form 298 (Rev. 2–89)
Prescribed by ANSI Std. 239–18

THIS PAGE INTENTIONALLY LEFT BLANK

QUALITY OF USMC OFFICERS: BUILDUP VS. REDUCTION IN FORCES

Michael S. Griner
Captain, United States Marine Corps
B.S., Southern Polytechnic State University, 2010

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
March 2016

Approved by: Marigee Bacolod
Thesis Advisor

Chad Seagren
Second Reader

William Hatch
Academic Associate,
Graduate School of Business and Public Policy

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

At the epicenter of maintaining the finest, most professional, and most feared military in the world is the quality of its leaders. The United States Marine Corps' Officer Corps is the standard bearer, with an accession pipeline more rigorous than any in the Department of Defense. As is the historical norm, however, the U.S. military undertakes ambitious increases to end strength in the wake of a prolonged conflict. As a conflict fades, the increased end strength is no longer warranted. The typical approach during the drawdown is to reduce accessions, create stricter retention policies, and entice members to leave the service through voluntary measures.

This research identifies the trade-off between quantity and quality necessitated by end-strength changes. Quantitative analysis using a difference-in-differences research design shows, relative to the buildup, officer quality increases during the drawdown. In particular, combat and non-combat occupations have a 0.0321 and 0.0834 point increase, respectively, in FITREP scores in the drawdown compared to the control group. Alternative measures and additional robustness checks support the hypothesis that the drawdown yields higher-quality officers. It is imperative to adapt force-shaping policies to a gradual approach to ensure the Marines retained during the buildup meet the quality standards of the drawdown.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AC	Active Component
ACT	American College Testing
ALMARS	All Marine Corps Activities
ASVAB	Armed Services Vocational Aptitude Battery
CBO	Congressional Budget Office
CD	Career Designation
CMC	Commandant of the Marine Corps
CNA	Center for Naval Analyses
DiD	Difference-in-Differences
DOPMA	Defense Officer Personnel Management Act
EAS	End of Active Service
FITREP	Fitness Report
FOIA	Freedom of Information Act
GCT	General Classification Test
GWOT	Global War on Terrorism
ID	Identification
MARADMIN	Marine Administrative Message
MBS	Master Brief Sheet
MCO	Marine Corps Order
MCRC	Marine Corps Recruiting Command
MCU	Marine Corps University
MMRP-30	Manpower Management Records and Performance-30
MOS	Military Occupational Specialty
MRO	Marine Reported On
PARS	Performance Anchored Rating Scale
PES	Performance Evaluation System
RIF	Reduction in Forces
RO	Reviewing Officer

RV	Relative Value
SAT	Scholastic Aptitude Test
SER	Selective Early Retirement
SSB	Special Separation Benefit
TBS	The Basic School
TERA	Temporary Early Retirement
USMC	United States Marines Corps
VSI	Voluntary Separation Incentive
WWII	World War II

ACKNOWLEDGMENTS

First and foremost, I would like to think my savior Jesus Christ. Without my faith I could have never fulfilled the rigorous requirements set forth by the Marine Corps. Secondly, I would like to thank “the wife,” Lindsey Griner (the wife is the term of endearment I most frequently use in reference to her). Her loyalty and dedication to me have never ceased, even though the Marine Corps journey has not been easy. She is truly the “better half” and without her support and care for our beautiful daughters, Bella Jean and Aubrey, I could have never completed this degree. A big thanks to my parents is definitely warranted; my upbringing was very grounded and their life lessons have guided me thus far.

A huge thanks must be given to my professor, advisor, and friend, Marigee Bacolod. Our tenures at NPS began at the same time. Her dedication to the students surpassed that of most professors, ensuring that someone with a limited mental capacity grasped the required knowledge. Her previous positions did not allow her the opportunity to work closely with the military, but she is a true professional and has learned and continues to learn about our organizations and how they function.

To Professor Seagren, thank you for your efforts in serving as my second reader and for your classroom instruction. Your lessons will truly benefit me in my future endeavors. I really appreciate the efforts of Tim Johnson at TFDW; thank you so much for providing the personnel data, more than once, to make this thesis possible. Your punctuality prevented the normal stress of data requests that can take several months. Doreen Marucci at MMRP-30, thank you for taking the time to provide the FITREP data during a very stressful time preparing for promotion boards. Without your efforts, my thesis would have no true basis of quality and the results would lack credibility.

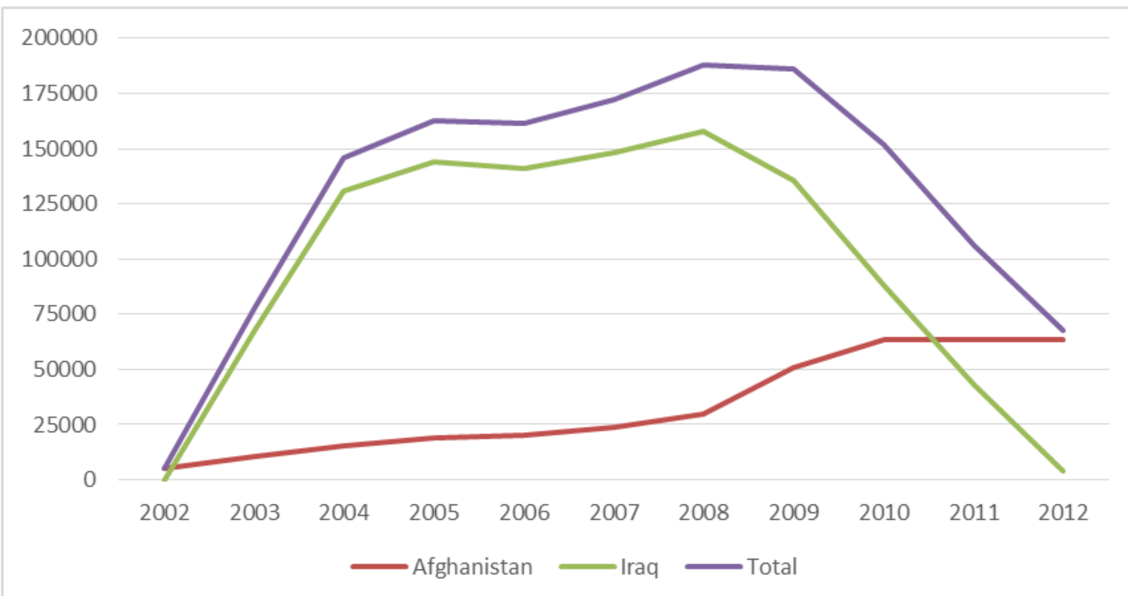
THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

In the aftermath of September 11, 2001, the United States declared a Global War on Terrorism (GWOT), which shifted military focus toward operations against organizations deemed “terrorist” as well as regimes accused of supporting them. President George W. Bush ordered strikes in Afghanistan in October 2001 and again in Iraq in March 2003 (“Timelines,” 2013). This two-front war required vast resources and manpower. Figure 1 displays the manpower levels throughout the wars summed across all U.S. military services. By 2013, President Obama declared the GWOT over (Shinkman, 2013).

Figure 1. Boots on the Ground Iraq and Afghanistan

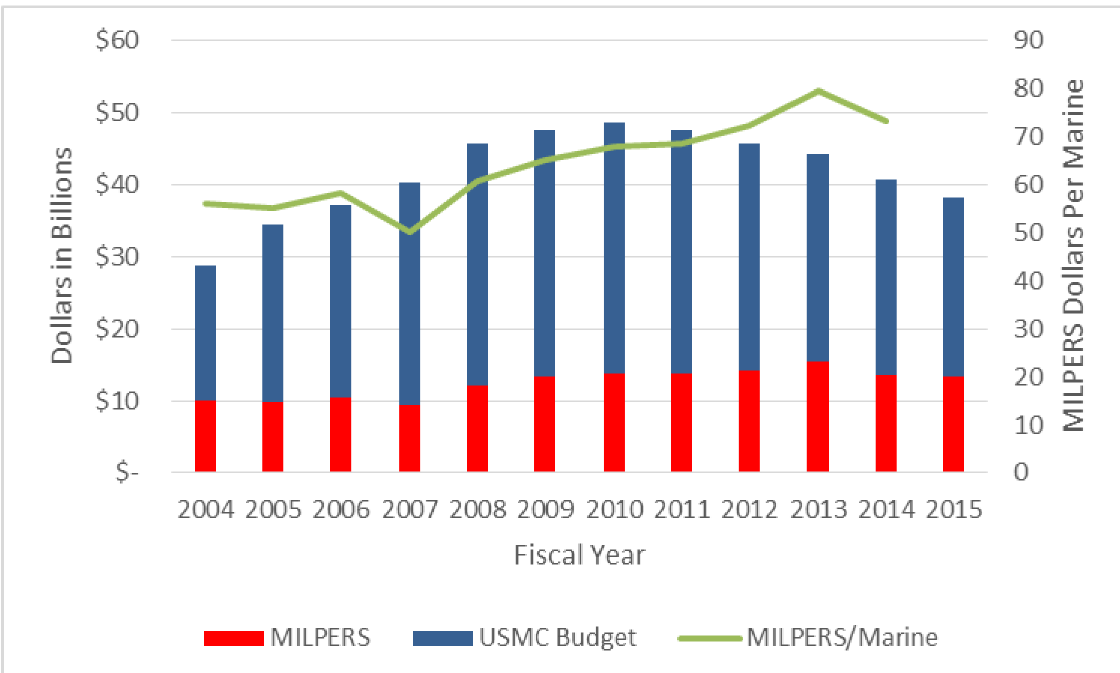


Adapted from Belasco, A. (2009). *Troop levels in the Afghan and Iraq wars, FY2001–FY2012: Cost and other potential issues* (CRS Report No. R40682). Retrieved from https://www.fas.org/sgp/crs/natsec/R40682.pdf

Building up and maintaining the military manpower required to fight a two-front war is not costless. Manpower expenditures typically comprise over 25 percent of the annual appropriations of the Department of Defense (DOD). Figure 2 portrays the costs

of military personnel (MILPERS) within the United States Marin Corps' (USMC) budget; interestingly, the MILPERS spent per Marine increases with the end strength. Given fiscal constraints, trade-offs between quantity and quality of troops were likely inevitable during the GWOT buildup and subsequent drawdown. This thesis is an empirical investigation of the quantity-quality trade-off among USMC officers during the buildup and drawdown.

Figure 2. USMC Military Personnel (MILPERS)



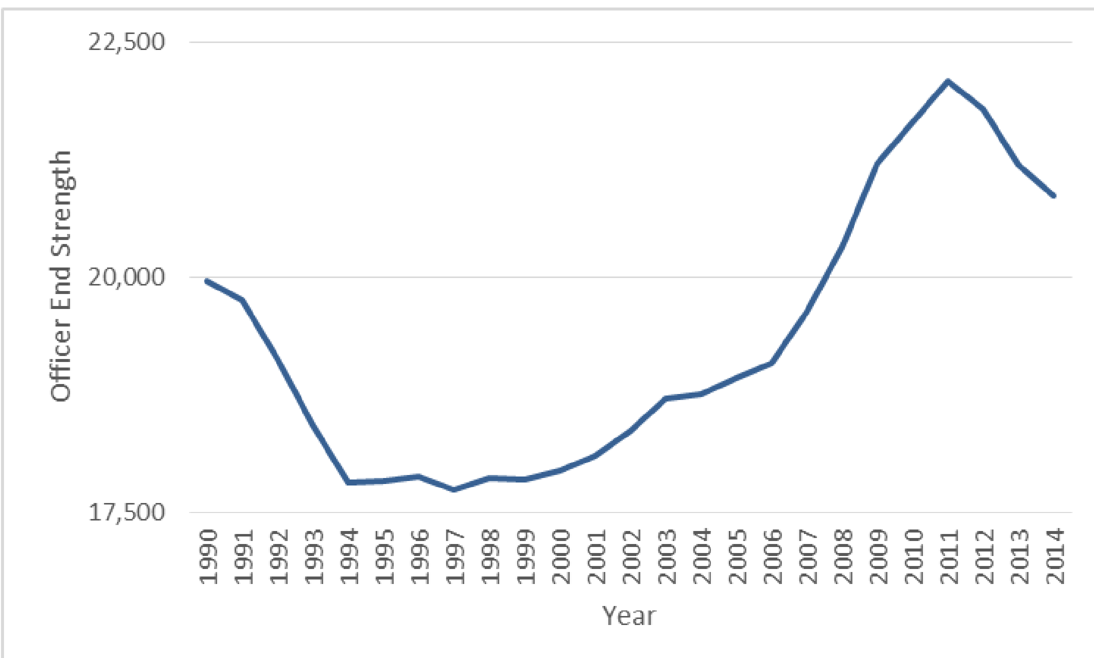
Adapted from Office of the Under Secretary of Defense (Comptroller). 2005, 2006, 2007, 2009, 2011, 2012, 2013, 2014, 2015. Defense budget materials FY2004–FY2016: Military personnel programs (M-1). Retrieved from http://comptroller.defense.gov/BudgetMaterials.aspx#summary and adapted from Budget of the U.S. Navy and the U.S. Marine Corps from fiscal year 2000 to 2016 (in billion U.S. dollars). (2016). Statista. Retrieved from http://www.statista.com/statistics/239290/budget-of-the-us-navy-and-the-us-marine-corps/

In February 2007, General James Conway, 34th Commandant of the Marine Corps (CMC), published All Marine Corps Activities 008/07 (USMC, 2008c), titled *Marine Corps End Strength Increase*. His intent was to inform all Marines that the president approved his plan to expand the USMC end strength from 175,000 Marines to 202,000. General Conway's plan to grow the force predated President Bush's decision to

deploy surge forces to Iraq (USMC, 2008c). With an extended two-front war looming, it was even more important to grow the force.

Growing any organization by more than 15 percent is a daunting task. Growing a professional military by 15 percent is even more difficult. The support required to train and equip a service member is extensive and costly. However, it is much easier to grow a force rapidly than it is to downsize the force in a quick, fair manner, and with the quality demanded by the military. The Corps' initial goal was to reach an end strength of 202,000 by fiscal year (FY) 2011; the goal was attained two years early (Sneden, 2009). At its peak, officer end strength was 22,094 (USMC History Division, 2014). Figure 3 depicts a chart of USMC officer end strength from 1990 to 2014. The 1990s represents the last occurrence of a drawdown, which happened during President Clinton's administration.

Figure 3. USMC Officer End Strength 1990–2014



Adapted from USMC History Division. (2014). Marine Corps end strengths. Retrieved from http://www.mcu.usmc.mil/historydivision/pages/frequently_requested/EndStrength.aspx

The USMC end strength was initially in flux once it was determined to drawdown the force. According to research conducted by the Sustainable Task Force, the Marine Corps could have been reduced to 175,000 Marines (Lamothe, 2010). Instead, the Pentagon determined that the Marine Corps could maintain an end strength of 184,000 for FY16 (Sisk, 2015). It appears that the consensus is to maintain the force near FY16 levels for the foreseeable future.

Numerous policy alternatives are available to reduce the size of the force. Reduction methods include reduced accessions, reduced retention, offering early retirement incentives, not honoring obligatory contracts with service members, and using any combination of the methods. This study is not intended to evaluate the methods, but more information can be found in the Center for Naval Analyses (CNA) study *Right-Sizing the Force* (Rostker, 2013). The Marine Corps decided to do a combination of all methods, except for not honoring contracts.

1. Career Designation

Title 10 U.S. Code requires that almost every abled officer be retained and promoted to the rank of captain. An officer is promoted to captain at four to five years of time in service and to major only after nine to 10 years of service. The USMC created the career designation (CD) board to retain the highest quality officers at the rank of captain. This creates a retention milestone sooner than the nine-to-ten-year mark corresponding with promotion to major. CD serves as another tool to shape the force outside of the gradual approach in Title 10.

Essentially, the CD board is a promotion-style board that reviews eligible officers and determines if their future service is in the best interest of the USMC. To be eligible for competitive CD, an officer must attain multiple prerequisites, but the four most pertinent as listed in MARADMIN 021/09 are:

1. Unrestricted officer in the active component.
2. 540-days of observed FITREP time. Officers can verify their observed FITREP time by contacting the Manpower Management Support Branch (MMSB).

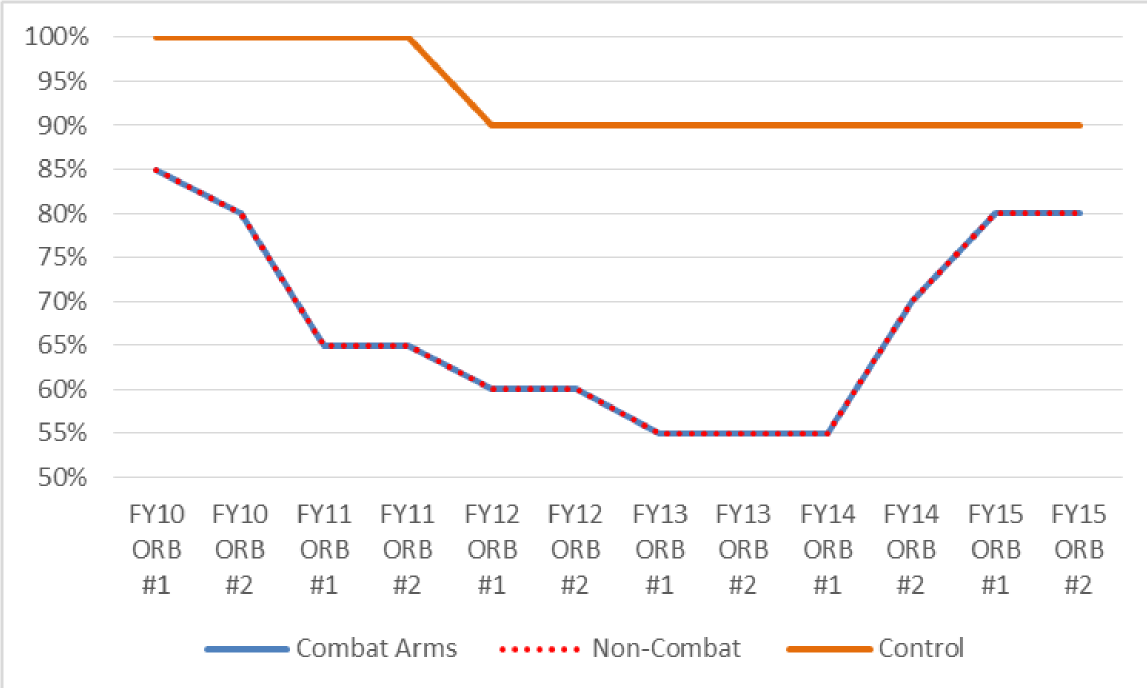
3. An officer's first opportunity for career designation consideration will coincide with that officer being in the promotion zone for the captain promotion selection board. Officers who do not meet the 540-day observed time requirement when in-zone for Captain will be considered for career designation at a subsequent board after achieving 540-days observed time.

4. Officers will be granted at least one opportunity to be considered for career designation. Officers with an Expiration of Active Service (EAS) that does not support at least one opportunity for career designation must request (via AA form) an extension on active duty to be considered. (USMC, 2010a)

Prior to the drawdown, CD was nearly a guarantee since it was not a competitive process. Fighting a two-front war required the USMC to retain every abled body to achieve its mission and to attain the congressionally mandated 202,000 end strength. General Conway, CMC, decided that returning to competitive CD would assist in reducing the size of the Corps. As stated in the prescribed Marine Administration Message 021/09 (USMC, 2010a), "Career Designation is a force-shaping tool that allows for the management of the officer population by retaining the best qualified officers from each year group. Those selected for Career Designation are offered the opportunity to remain on active duty" (USMC, 2010a). Essentially, competitive CD is a method to prescriptively shape the force outside of mandated Title 10 requirements.

Figure 4 is derived from CD board results from FY10 to FY15. For CD boards, Military Occupational Specialties (MOS) are normally placed into five competitive categories: combat arms, combat service support, aviation support, aviation, and law. Given this study's research design (discussed later in more detail), these groups were merged into three. The group referred to as "control" is the combination of aviation and law. The first experimental or "treatment" group is combat arms. The second is the non-combat group and consists of combat service support and aviation support. Figure 4 shows that CD selection was highly likely for these groups around 2010. With the drawdown from an officer end strength over 22,000 and competitive CD, the combat and non-combat groups reached a low CD percentage of 55 percent in FY13–FY14, as seen in Figure 4.

Figure 4. CD by MOS Groups



Adapted from Johnson, J. A. (2015). *Significant pre-accession factors predicting success or failure during a Marine Corps officer's initial service obligation.* Manuscript submitted for publication at Naval Postgraduate School.

The CD board and captain promotion board are not an integrated process. The CD board is conducted semi-annually in February and August, while the captain promotion board is conducted annually in February. Title 10 U.S. Code requires that all fully qualified first lieutenants eligible for promotion shall be promoted to captain. Since captains are required to be promoted 100 percent by law it becomes very difficult to quickly reduce the officer corps. Figure 5 displays the DOD Instruction 1320.13 (DOD, 2014) that enforces Title 10 and shows the desired timing and opportunity for promotion into the ranks of major, lieutenant colonel, and colonel. CD benefits the Marine Corps because it allows the number of officers to be reduced before they reach promotion zone for major.

Figure 5. Field Grade Promotion Timing and Opportunity

DESIRED ACTIVE-DUTY LIST PROMOTION TIMING AND OPPORTUNITY		
TO GRADE	TIMING[1,2]	OPPORTUNITY[1,3]
O-4	10 Years +/– 1 Year	80 Percent
O-5	16 Years +/– 1 Year	70 Percent
O-6	22 Years +/– 1 Year	50 Percent
1. Promotion timing and opportunity are defined in Reference (e). 2. Years of commissioned military service plus all entry grade credit. 3. Promotion opportunities will comply with Reference (c), section 623(b)(4).		

Source: Department of Defense. (2014, October 30). *Commissioned Officer Promotion Reports* (DOD Instruction 1320.13). Retrieved from http://www.dtic.mil/whs/directives/corres/pdf/132013p.pdf

2. Fitness Reports

The FITREP is the most important evaluation tool for Marine Corps officers. "The FITREP provides the primary means for evaluating a Marine's performance to support the Commandant's efforts to select the best qualified personnel for promotion, career designation, retention, resident schooling, command, and duty assignments" (USMC, 2015, p. 2). To meet career milestones it is imperative that a Marine's FITREPSs are current, accurate, and most importantly competitive relative to his/her peers. FITREPS are important for all Marines, but even more so for young officers. As previously stated, the CD timeframe is very condensed and occurs early in an officers' career. Second and first lieutenants are the only ranks that receive semi-annual FITREPSs. Semi-annual reports allow the 540 days of observation to be computed twice per year, coinciding with the two CD boards. Semi-annual reports also prevent population overload on any single CD board. This allows the force to more accurately shape the distribution of officers.

The FITREP evaluates a Marine on 13 attributes that are categorized into four sections: mission accomplishment, individual character, leadership, intellect and wisdom (MCO 1610.7, p. 4-21). It is designed to evaluate a Marines' work performance, while also accounting for off duty actions. The reporting senior (RS) is defined as: "the first commissioned or warrant officer (or civilian GS-9/equivalent or above) in the reporting chain senior to the MRO" (USMC, 2015, p. 2-1). The FITREP is a performance anchored

rating scale (PARS) that assigns a numeric value to each selected trait. As stated in MCO 1610.7, the numeric grading is as follows:

> Each block in the marking gradient for each attribute has an assigned numeric value as follows: A=1, B=2, C=3, D=4, E=5, F=6, G=7, and H=0 (not observed). NOTE: Block H (not observed) has no value and does not factor into the calculation of the average. Each observed attribute numeric value is added to find the total, which is divided by the number of observed attributes to calculate the average, rounded to the nearest hundredth. (USMC, 2015, pp. 8-4–8-5)

Each report receives an average of all rankings and is the first time a numeric value is assigned to a report. The Marine reported on (MRO) can view his/her master brief sheet (MBS) and view the average of their report and compare it to the average of the reports written by that RS in the same grade. The MRO can also see the highest score written by the RS and use it as a reference for their score. Once the RS has written three observed FITREPSs on Marines of the same grade, each report will be assigned a relative value (RV) between 80 and 100 percent.

> The relative value will appear on the MRO's MBS in numeric fashion on an 80 to 100 scale. (1) A relative value between 93.34 and 100.00 indicates the report is in the upper third of the RS profile. (2) A relative value between 86.67 and 93.33 indicates the report is in the middle third of the RS profile. (3) A relative value between 80.00 and 86.66 indicates the report is in the bottom third of the RS profile. (USMC, 2015, p. 8-6)

The reviewing officer (RO) is the direct supervisor of the RS. The RO must "ensure the integrity of the system by giving close attention to accurate marking, narrative assessment, and timely reporting" (USMC, 2015, p. 2-2). The RO can concur or not with the RS's assessment of the MRO. The RO also ranks the MRO according to their assessment. The RO's ranking is converted numerically and the MRO can compare their score with other MROs of the same rank evaluated by the RO. This study does not use the RO's average rankings to draw any conclusions; it focuses solely on the RS averages at the time of reporting.

B. MOTIVATION: TRADE-OFF BETWEEN QUANTITY AND QUALITY

Buildups and drawdowns are necessary to maintain a professional military force. Throughout American history, buildups occur at the start of a conflict and occur rapidly at the conclusion of the conflict. Issues then arise with maintaining the quality of the force with these rapid changes in quantity. During the buildup the initial policy is to retain everyone quickly to reach the congressionally mandated end strength and quality is likely neglected in the process. Once the conflict ends, the force must return to its pre-war size or what Congress determines should be the new end strength to stay prepared for future adversaries.

The subsequent drawdown requires another policy and usually results in lower accession, early retirement opportunities, and stringent retention policies. Company grade officers become the easiest population with which to establish strict retention policies. The CD process becomes much more competitive, yet the field grade ranks are populated with officers retained during the noncompetitive buildup. This creates field grade ranks populated with officers of all types of quality, whereas the most competitive CD periods tends to select above-average officers due to strict retention policies.

It thus becomes challenging for the USMC to systematically reduce the officer population during the drawdown while maintaining some desired level of quality. Incentives must be created to entice individuals to leave or the Corps has to force them out. Another problem is ensuring that low quality officers take the incentives to leave rather than high-quality officers. It is a very lengthy process to naturally purge the system without immediate incentives. Meanwhile, Title 10 regulates all officer promotions.

C. PURPOSE

The purpose of this study is to evaluate the effect or net change in the quality of the USMC officer population during the buildup and subsequent drawdown of forces. The natural hypothesis is that the average quality of USMC officers is lower when retaining nearly all officers during the buildup, and the implementation of competitive CD would increase average quality. This study forms empirical estimates of these trade-offs in quantity and quality of the officer corps.

D. ASSUMPTIONS

The primary assumption for this research is that the FITREP, as an output based measure of quality, is the best measure of officer quality available. Other individual input-based measures of quality are the physical fitness test (PFT), swim qualifications, rifle scores, the combat fitness test (CFT), and so forth. These input measures are used to predict career designation, but cannot solely be used to determine the overall quality of an officer. From a production function perspective, these input measures may also display diminishing returns or contributions to an officer's quality, productivity, job performance, or output. The FITREP is designed for the RS to take into account the individual tests, physical and combat standards, as well as leadership qualities. Therefore, the FITREP is the best, most comprehensive measure of quality for an individual officer.

To accept the FITREP as the best measure of quality, it must be assumed that the evaluations are not inflated and that they accurately measure an officer's performance or quality. According to MCO 1610.7, "Countering inflation begins with the reporting officials, specifically the RS and RO, who must accurately report a Marine's performance." To reduce inflation tendencies, MCO 1610.7 states:

> 1. The design of this report limits the ability of RSs to unjustifiably inflate a Marine's performance.
>
> 2. Reports must be based on a Marine's performance vice sociability. Reporting officials can inadvertently render these controls ineffective by preparing and submitting FITREPSs that fail to adhere to both the letter and the spirit of this Manual. (USMC, 2015, p. 1–2)

E. RESEARCH QUESTIONS

The primary research question for this thesis asks: During the buildup and subsequent drawdown of forces between 2007 and 2013, what was the net change in the quality of USMC officers, as measured by the probability of CD selection and job performance measures?

Secondary questions include: Are Cancian and Klein (2015) correct in their assessment that the quality of USMC officers declined? Cancian and Klein measure quality using performance on the General Classification Test (GCT) and examined the

years 1980 to 2014. Is the GCT an appropriate measure of officer quality and do alternative measures mirror GCT changes during this period?

THIS PAGE INTENTIONALLY LEFT BLANK

II. LITERATURE REVIEW

This research is focuses on the net change in quality of USMC officers that occurred as a result of the buildup and drawdown of forces at the conclusion of a two-front war. The studies reviewed in this chapter assess different variables that affect the quality of a force and also the mechanisms to reduce the force in an efficient fair manner.

A. CANCIAN AND KLEIN

Cancian and Klein's research (2015) shows a decline in the quality of Marine Corps officers between 1980 and 2014. USMC officer performance on the General Classification Test (GCT) is the measure of quality used in the study. The GCT was developed during World War II and was intended to determine the learning abilities of incoming service members to better place them in the correct Marine Occupational Specialty (MOS). According to individuals' results on the GCT, they would be classified into categories ranging from one to five, five being the category for members with less potential to learn (as compared to soldiers with average mental ability), and one being the category for recruits with the highest ability (Cancian & Klein, 2015, p. 5). This test allowed the services to better distribute the influx of forces during the buildup period.

The test was designed to have an average grade of 100 with a standard deviation of 20 (Cancian & Klein, 2015, p. 5). The GCT was nearly phased out in 1973 in favor of the Armed Services Vocational Aptitude Battery (ASVAB). The ASVAB was deemed the better test as Vietnam era conscription ended and the all-volunteer force began in 1973 (Cancian & Klein, 2015, p. 1). Using the GCT, Cancian and Klein note, that "while 85 percent of those taking the test in 1980 exceeded 120, the cut-off score for Marine officers in World War II, only 59 percent exceeded that score in 2014. At the upper end of the distribution, 4.9 percent of those taking the test scored above 150 in 1980 compared to 0.7 percent in 2014" (Cancian & Klein, 2015, p. 3).

Currently, the GCT is only administered by the Marine Corps to newly commissioned second lieutenants attending The Basic School (TBS). TBS is a six-month

training evolution that is required for all Marine Corps officers and is guided by the following mission statement:

> Train and educate newly commissioned or appointed officers in the high standards of professional knowledge, esprit-de-corps, and leadership to prepare them for duty as company grade officers in the operating forces, with particular emphasis on the duties, responsibilities, and warfighting skills required of a rifle platoon commander. (United States Marine Corps, n.d.)

During TBS an officer is evaluated on his or her intellectual, leadership, and physical fitness skills, with the highest priority given to leadership. The compilation of an officer's grades determines the probability that an officer gets the MOS that he or she desires. The GCT has no effect on this process, nor is performance on the GCT used to determine grades or an officer's subsequent career. In fact, anecdotal evidence and my own experience was that the GCT was not administered or, if the GCT was administered, officer candidates at TBS knew their GCT performance was of no particular consequence. As such, it is difficult to argue that GCT scores measure officer quality.

On the other hand, performance on the Scholastic Aptitude Test (SAT) is more commonly observed among officers, and the general population easily understands the scores. Cancian and Klein did obtain information from the Marine Corps Recruiting Command (MCRC) stating that from 2005–2014 SAT scores for new Marine officers averaged 1,200 points. In comparison, the average for high-school seniors that planned to attend college was 1,010 points (Cancian & Klein, 2015, pp. 9–10). The SAT is, arguably, the more relevant test, since it is easily understood and designed to adapt to generational gaps more than the antiquated GCT. They hypothesize that Marine officer quality declined, but they had only limited data to confirm the hypothesis. Furthermore, neither the GCT nor the SAT measures leadership qualities, which is arguably the most important attribute for a military officer. As such, evidence of a decline in the GCT and/or SAT performance of Marine officers may not necessarily reflect a reduction in quality.

In addition, Cancian and Klein's research analyzes data from 1980 to 2014, whereas this study focuses on the more recent officer cohorts between 2004 to 2015. This

research can only re-examine the overlapping years of 2004 to 2014 to verify Cancian and Klein's results. It would also not be prudent for me to analyze FITREPs dating back to 1980 due to many changes in the performance evaluation system and the rating instrument. It is nevertheless important to analyze these overlapping years, because it represents the current officer population. In contrast to Cancian and Klein's work, however, this research uses the GCT and SAT scores to determine if the tests have a statistically significant impact on the probability of being career designated. As stated in the introduction, this research's measure of quality relies not only on academic scores (input-based measure of quality), but also on FITREPs (output-based measure of quality). In their 2015 study, Cancian and Klein did not have access to FITREP information, since it is not available through a Freedom of Information Act (FOIA) request.

B. GARZA

Garza's research (2014) identifies the significant predictors that could be used to determine the competitiveness of a company grade officer for the career designation (CD) board. It is very important to note that the competitive CD process did not start until FY10 and data observed came from FY10–FY13. This research is different in that the concern is not only on what it takes to earn CD, but also that CD is viewed as a one-time screening process to determine future quality at a single point in time. This research evaluates the average quality of the force as measured by changes in FITREPs and not by individual predictors of success for CD.

C. THE DRAWDOWN OF THE MILITARY OFFICER CORPS

The U.S. military end strength increased during the Cold War. Between 1989 and 1996, post-Cold War, the force needed to drawdown its forces (CBO, 1999, preface). This is very similar to the current situation, as operations in Iraq have concluded and operations in Afghanistan are in the final stages of drawdown. To say that history is repetitive is an understatement. The post-Cold War Congressional Budget Office (CBO) report states, "The managers faced a difficult challenge: to bring in enough new officers to maintain a combat-ready force in the future and yet keep faith with personnel already in uniform" (CBO, 1999, p. 3). These very same statements are echoed throughout the

Marine Corps during the current drawdown period. The 35th CMC, General James Amos, stated that, "while we are working hard to balance our myriad requirements, I want each of you to know that keeping faith with you and your families is a top priority of mine—I consider it a sacred responsibility" (deGrandpre, 2013).

Reduction in forces (RIF) aims for as many service members as possible to voluntarily leave the service on self-imposed terms. This prevents difficult decisions that may separate a quality officer, mid-career, when attaining the 20 years to guarantee a pension were well within reach. Voluntary separations can be accomplished by many means. The methods used in the Obama administration mirror those used in the Clinton administration. The plans may have minute differences, but for the purpose of this study, they have the same effect. The tools at the services' disposal are voluntary separation incentive (VSI), special separation benefit (SSB), and temporary early retirement (TERA). Another tool is selective early retirement (SER); however, if selected it forces an officer to retire regardless of the limitations outlined in the Defense Officer Personnel Management Act (DOPMA).

This research does not focus on members leaving the service due to early separation; however, it is important on two fronts:

1. The fact that the buildup kept almost all officers means that there was likely an impact on the average quality of the force.
2. A follow-on study should be conducted to determine if these programs targeted the correct population that would retain the highest quality officers, while separating the lower quality. Policy revision should be warranted if the wrong population was targeted for reduction.

III. DATA AND METHODOLOGY

This chapter describes the various entities that provided data, the numerous variables used, and the methodology for data analysis. All data spans the timeframe of FY04–FY15.

A. DATA SOURCES

This research provides a framework to analyze the net change in quality of Marine Corps officers that occurred as a result of the buildup to an end strength of 202,000 and the subsequent reduction in forces to 184,000 (Sisk, 2015). Personnel performance data and FITREPSs are needed to accurately determine and analyze any changes in quality. The following section lists and describes the data sources used for the subsequent analyses.

1. Total Forces Data Warehouse (TFDW)

All officers' characteristics, demographics, and training information for this research is provided by Total Forces Data Warehouse (TFDW), which is defined as

> a system of the Manpower Information Technology Branch of Manpower & Reserve Affairs (M&RA). It is the Marine Corps' official system of record for USC Title 10 end strength reporting. The TFDW houses more than 30 years of historical manpower data from a variety of USMC and DOD systems including MCTFS, MASS, RCCPDS, MCTIMS and DEERS, in one central location to provide manpower analysts with a comprehensive view of a Marine's career from "street to fleet." (TFDW, n.d.)

Data from TFDW consist of eight, non-symmetrical, personnel databases that were stripped of personally identifiable information (PII) prior to transmission. TFDW provided a randomly-generated unique identifier for each personnel, so it could later be merged with the FITREP database. However, the personnel files from TFDW contained many observations with the same identifier, as an officer's record was updated within the same year. To ensure the data contain only one record per officer-FY, and that this data

reflect characteristics of a Marine officer (e.g., marital status) at the end of a fiscal year, each of the eight files had to be cleaned before being merged.

2. Manpower Management Records and Performance-30 (MMRP-30)

MMRP-30 is the performance evaluation section of the Manpower Management Division within Manpower and Reserve Affairs (M&RA). MMRP-30's mission is "to support the Marines' needs in ensuring their performance evaluation records are up to date and prepared for promotion boards" (MMRP-30, n.d.). MMRP-30 maintains all FITREP data for the Marine Corps. Similar to TFDW data, all PII information was stripped, and officers are identified only through a randomly-generated identifier. The data from MMRP-30 came in seven different files; however, the data originated from the same database, it was stored in the same manner, and the data was easily appended.

3. Sample Restriction and Number of Observations

Once cleaned, limited to the correct officer population, and merged using the unique identifier, the data represent 172,771 officer-FY observations for analysis. The observations span FY04–FY15 and are restricted to USMC officers at the ranks of first lieutenant to lieutenant colonel. A full list of descriptive statistics are available in Appendix E.

B. VARIABLES

1. Dependent Variables

a. Career Designation (CD)

CD is a force-shaping tool that intends to retain the Corps' desired level of quality officers. The requirements for CD eligibility can be found in MCO 1001.45J and in the introduction of this thesis. CD is a binary variable that takes on a value of one if an officer is offered and selects CD or is offered CD and declines the offer. Declining implies the officer possessed the qualities desired by the Corps, but the individual voluntarily executed their end of active service (EAS) obligation. All other values of CD are coded as zero.

It is also useful to note that given the sample restriction, the research focuses on the point in time when non-CD, changes to CD in an officer's career, or when the officer's status never changes and he or she subsequently exits the USMC active component. Selection to Captain is not viewed as a measure of quality since Title 10, U.S. Code requires all eligible officers to be promoted to the rank of captain.

b. FY Average of FITREPSs

As previously explained, a FITREP attains an average rating based on the marks given by the RS. A Marine officer is never stagnant in any one position. Company grade officers frequently change jobs to gain experience and demonstrate versatility. Usually, every positional change coincides with an occasion that mandates a new FITREP. Reporting occasions can be found in MCO 1610.7, the Performance Evaluation System (PES). It is not uncommon for an MRO to receive multiple occasional reports throughout any given FY.

This research's goal is to analyze net quality changes based on FY. If an officer has multiple reports during a FY, the data first has to be manipulated to create an FY average. To do so, the data is transformed wide by FY and identification (ID) number and subsequently the report averages for that FY were averaged. This created the variable "avg_this_rpt"; which is the average of the report averages for a FY. The data is then retransformed to long data, thus creating only one average for each FY. This maintains the integrity of the data and prevents dropping any observations for a given officer. Once transformed, this allows "avg_this_rpt" to be a dependent variable for analyses.

c. FY Average of FITREPSs minus FY Average of Average RS High

Additionally, it is interesting to analyze any change in the dispersion of each report average relative to the report highs of a given RS. Once again there may be multiple reports for an MRO in a given FY. The data was transformed wide and a variable was created for FY average of the average RS high ratings. The data was then transformed long so that there is only one average for each officer during the FY.

A new variable called "avg_hi" is created to reflect the FY FITREP average minus the average of the RS high reports. Thus, "avg_hi" measures the deviation of an officer's ratings from the RS' historically highest rating. The more negative this variable is, the lower the officer's quality relative to his or her peers in the given period. Vice versa, the closer to zero or even positive this variable becomes, the higher the quality on average during this period.

2. Independent Variables

Only those independent variables that are not easily understood are discussed in this section. Ethnicity/race is not included as a control variable in this study for two reasons:

1. The study is not concerned with analyzing quality across ethnicities/race, although it is recommended as follow-on research.
2. The databases from TFDW containing ethnicity/race appeared unreliable and difficult to reconcile across personnel files, and a lot of assumptions would have to be made to identify an officer's ethnicity/race. Therefore, the researcher determined adding race to the model would only introduce statistical noise.

This research is not concerned with the probability of CD based upon individual military occupation specialties (MOSs). However, the research design, described in more detail in this section, necessitates grouping the MOSs and analyzing several years of data by MOS groups. In general, lawyers and pilots are accepted for CD at rates over 95 percent; therefore, I use these two specialties as the control group in an experimental research design. The remaining MOSs are divided into combat and non-combat groups. All MOS groups are coded as binary variables that take the value of zero if an officer's MOS does not belong to that group and one if it does. The three MOS classification groups are thus the control group, for whom end strength policy changes have little to no impact, and the two "treated" groups: combat, and non-combat.

3. Performance Independent Variables

The data contains numerous variables for individual performance tests that were military requirements or that may have been taken outside of the service and entered in

one's record. The American College Testing (ACT) and Scholastic Assessment Test (SAT) are standardized tests usually taken by high school students in preparation for college applications. Both tests are used to for college acceptance and admittance requirements.

The USMC requires an SAT of at least 1,000 to be commissioned an officer. Recently, the SAT was revised changing the maximum score to 2,400 from 1,600. It was assumed that all officers in this data took the older SAT with a maximum score of 1,600. The data contained 11 scores greater than 1600, and those scores were truncated to 1,600. There were 691 scores below 1,000, and those scores were edited to the required 1,000. However, there are 154,155 scores that are zero or missing in the data. A missing SAT variable (xsat) was created to indicate this missing category and prevent deleting 154,155 observations from the analysis. The effect of missing SAT on officer quality can now also be analyzed.

The Law School Admission Test (LSAT) is included in the data and 88 percent of those scores fall within the MOS control group containing lawyers. This is obvious since the control group population represents lawyers and pilots. This research used rifle range score classes to analyze the effect on specific dependent variables. The rifle range classes were coded as zero for all classes that were not expert and one for expert class.

Regional, Cultural, and Language Familiarization (RCLF) is a fairly new initiative by the USMC. Each new officer at TBS is assigned a different region to become familiar with throughout their career. The familiarization process requires multiple, usually online courses, to be completed at each rank. The mission of RCLF is: "to ensure that Marine units are globally prepared and regionally focused so they are effective at navigating and influencing the culturally complex 21st Century operating environment in support of the Marine Corps' missions and requirements" (Marine Corps University, ,n.d.). RCLF was not used in this research due to only a small portion of the drawdown population being subjected to RCLF. An even smaller number of officers have completed RCLF due to lack of time in the system with the new requirement.

Swim qualifications underwent a revision during the time period of the data. Marine Corps Water Survival Training Program (MCWSTP), MCO 1500.52D, mandates the USMC swim qualification requirements (USMC, 2010b). For the purposes of this research swim qualification was coded as a binary variable. Officers with a qualification of “Q,” “1,” “2,” “Advanced,” “Marine Corps Instructor of Water Survival (MCIWS),” and “Marine Corps Instructor Trainer of Water Survival (MCITWS)” were coded with a one and all others were coded with 0.

The physical fitness test (PFT), as prescribed in MCO 6100.13 differentiates classes of PFT scores into three main categories. The categories are first, second, and third classes, with first class corresponding to a higher score and better physical fitness. The numeric scores can range from 0–300, with 300 being a perfect score. To reduce mitigating effects this research placed PFT classes into two categories. The variable “pft_class” is a binary variable that is assigned the value of one if a first class PFT is achieved and zero otherwise. A binary variable called “weight_control” was created to identify an officer that has ever been assigned to weight control. The variable is assigned the value of one if the officer has ever assigned and zero otherwise.

C. METHODOLOGY

The research design used in this research is a difference-in-differences (DiD) analysis, a method of program evaluation typically used by applied economists. In the absence of a social experiment where a policy intervention, or in the language of econometrics, the treatment, is randomly allocated across the population, the DiD design mimics an experimental research design using observational data.

One way to estimate the impact of end strength policy changes is to simply examine outcomes before and after the policy change. This is referred to as a simple Difference estimator in econometrics. Within the subject population (USMC officers), what was the average change in quality between 2004 and 2015? This would require estimating models of the form:

$$Y_{it} = \beta_0 + \beta_1 T + \varepsilon_{it} \tag{1}$$

where Y measures the outcome or quality measure, and T=1 if the Officer *i* is observed after the policy change (Gertler et al., 2011, pp. 95–98). However, one problem with the simple Difference estimator is that we cannot distinguish if this change over time in the outcome *Y* is due to the policy change or from secular changes (for instance, if USMC officer quality is on an upward trend due to better recruiting practices or worsening U.S. economy).

To improve on the simple Difference estimator and to estimate a causal effect of end strength policy changes, I compare the average change in outcomes over time between a group affected by the change (the treatment group) to a group I argue to be unaffected by the change (the control group). I designate the combat and non-combat groups to be treatment groups, and lawyers and pilots to the control group.

"The idea is to correct the simple Difference before and after for the treatment group by subtracting the simple Difference for the control group" (Duflo, n.d., p. 14). In the next section, I present the DiD regression models I estimate. In implementing the DiD method, the research eliminates or at the least mitigates biases such as mean reversion bias (the tendency for data to revert to the mean over time) and selection bias (as long as the selection is common to both treatment and control groups) (Gertler et al., 2011, pp. 95–98).

The crucial assumption for the DiD estimator to be an unbiased causal estimate of the impact of end strength changes on USMC officer quality is that, without a policy change, the average change over time in quality would have been the same for treatment and control groups (Duflo, n.d., p.13). In the language of econometrics, this is called the parallel trends assumption. The research examines the validity of this assumption in the robustness checks section of Chapter IV.

D. EMPIRICAL MODELS

1. Model 1: Probit CD during the Buildup FY07–FY09

Model 1, as shown in equation 2, is a probit regression model that analyzes the probability of being selected for CD during the buildup period, FY07–FY09.

$$\Pr\left(cd_{it} = 1\right) = \Phi\left(\beta_0 + \beta_1 buildup_t + \beta_2 M_i + \beta_3 buildup_t * M_i + \beta_4 perf_{it} + \beta_5 dem_{it}\right) \tag{2}$$

where buildup=1 if Officer *i* is observed during FY07-FY09 and zero otherwise; M is a vector of two indicator variables representing combat MOS and another for non-combat MOS. Thus, β_3 represents the DiD estimate of the net change in CD selection rates during the build-up period. It is the before-after change in CD between the treated groups (combat and non-combat) relative to the control group.

The performance (perf) variable is a vector of independent variables including: buildup interacted with the fiscal year (FY) average of the averages of each report written on a given MRO in a FY (buildup*avg_this_rpt), buildup interacted with the average report average minus the average RS report high for a given rank (buildup*avg_hi), ACT Score (act), Missing ACT Score (xact), GCT Score (gct), LSAT Score (lsat), Missing LSAT(xlsat), SAT Score (sat), Missing SAT (_sat), SAT interacted with Missing SAT (sat_sat), Class of Rifle Score (rifle_class=1 if first class, zero otherwise), AFQT Score (afqt), weight control status (weight_control=1 if an officer was ever assigned to weight control), PFT class (pft_class=1 if pft class is first class, zero otherwise, swim qual class (swim qual=1 if advanced, zero otherwise), and missing swim class (xswim). The demographic (dem) variable consists of: Marital Status (marital_status), Number of Dependents (num_dep), and Sex (sex, male=1, zero otherwise).

2. Model 2: Probit CD during the Drawdown FY10–FY13

Model 2 is a probit model that estimates the probability that an officer would be accepted for CD the drawdown period. The model is shown in equation 3:

$$\Pr\left(cd_{it} = 1\right) = \Phi\left(\beta_0 + \beta_1 drawdown_t + \beta_2 M_i + \beta_3 drawdown_t * M_i + \beta_4 perf_{it} + \beta_5 dem_{it}\right) \tag{3}$$

where drawdown=1 if Officer *i* is observed during FY10-FY13 and zero otherwise; M is a vector of two indicator variables representing combat MOS and another for non-combat MOS. Thus, β_3 represents the DiD estimate of the net change in CD selection rates during the drawdown period. It is the before-after change in CD between the treated groups (combat and non-combat) relative to the control group.

The performance (perf) variable is a vector of independent variables including: drawdown interacted with the fiscal year (FY) average of the averages of each report written on a given MRO in a FY (drawdown*avg_this_rpt), drawdown interacted with the average report average minus the average RS report high for a given rank (drawdown*avg_hi), ACT Score (act), Missing ACT Score (xact), GCT Score (gct), LSAT Score (lsat), Missing LSAT(xlsat), SAT Score (sat), Missing SAT (_sat), SAT interacted with Missing SAT (sat_sat), Class of Rifle Score (rifle_class=1 if first class, zero otherwise), AFQT Score (afqt), weight control status (weight_control=1 if an officer was ever assigned to weight control), PFT class (pft_class=1 if pft class is first class, zero otherwise, swim qual class (swim qual=1 if advanced, zero otherwise), and missing swim class (xswim). The demographic (dem) variable consists of: Marital Status (marital_status), Number of Dependents (num_dep), and Sex (sex, male=1, zero otherwise).

3. Model 3: FY Average FITREPs during the Buildup FY07–FY09

Model 3 is a linear regression model estimating the effects of the buildup on the FY average of the FITREP averages for a given MRO. Equation 4 represents the model.

$$avg\_this\_rpt_{it} = \beta_0 + \beta_1 buildup_t + \beta_2 M_i + \beta_3 buildup_t * M_i + \beta_4 perf_{it} + \beta_5 dem_{it} + \in_{it} \qquad (4)$$

where buildup=1 if Officer *i* is observed during FY07-FY09 and zero otherwise; M is a vector of two indicator variables representing combat MOS and another for non-combat MOS. Thus, β_3 represents the DiD estimate of the net change in job performance of USMC officers during the build-up period, as measured by FITREP averages. It is the before-after change in average job performance between the treated groups (combat and non-combat) relative to the control group.

The performance (perf) variable is a vector of independent variables including: ACT Score (act), Missing ACT Score (xact), GCT Score (gct), LSAT Score (lsat), Missing LSAT(xlsat), SAT Score (sat), Missing SAT (_sat), SAT interacted with Missing SAT (sat_sat), Class of Rifle Score (rifle_class=1 if first class, zero otherwise), AFQT Score (afqt), weight control status (weight_control=1 if an officer was ever assigned to weight control), PFT class (pft_class=1 if pft class is first class, zero otherwise, swim

qual class (swim qual=1 if advanced, zero otherwise), binary variable for CD (cd=1 if accepted, zero otherwise), and missing swim class (xswim). The demographic (dem) variable consists of: Marital Status (marital_status), Number of Dependents (num_dep), and Sex (sex, male=1, zero otherwise).

4. Model 4: FY Average FITREPs during the Drawdown FY10–FY13

Model 4 is a linear regression model estimating the effects of the drawdown on the FY average of the FITREP averages for a given MRO. Equation 5 represents the model.

$$avg\_this\_rpt_{it} = \beta_0 + \beta_1 drawdown_t + \beta_2 M_i + \beta_3 drawdown_t * M_i + \beta_4 perf_{it} + \beta_5 dem_{it} + \in_{it} \quad (5)$$

where drawdown=1 if Officer *i* is observed during FY10-FY13 and zero otherwise; M is a vector of two indicator variables representing combat MOS and another for non-combat MOS. Thus, β_3 represents the DiD estimate of the net change in job performance of USMC officers during the drawdown period, as measured by. It is the before-after change in average job performance between the treated groups (combat and non-combat) relative to the control group.

The performance (perf) variable is a vector of independent variables including: ACT Score (act), Missing ACT Score (xact), GCT Score (gct), LSAT Score (lsat), Missing LSAT(xlsat), SAT Score (sat), Missing SAT (_sat), SAT interacted with Missing SAT (sat_sat), Class of Rifle Score (rifle_class=1 if first class, zero otherwise), AFQT Score (afqt), weight control status (weight_control=1 if an officer was ever assigned to weight control), PFT class (pft_class=1 if pft class is first class, zero otherwise, swim qual class (swim qual=1 if advanced, zero otherwise), binary variable for CD (cd=1 if accepted, zero otherwise), and missing swim class (xswim). The demographic (dem) variable consists of: Marital Status (marital_status), Number of Dependents (num_dep), and Sex (sex, male=1, zero otherwise).

5. Model 5: FY Average FITREPSs minus FY Average RS High during the Buildup FY07–FY09

Model 5 is a linear regression model estimating the effects of the buildup on the FY average of the FITREP averages minus the FY average RS high for a given MRO. Equation 6 represents the model.

$$avg\_hi_{it} = \beta_0 + \beta_1 buildup_t + \beta_2 M_i + \beta_3 buildup_t * M_i + \beta_4 perf_{it} + \beta_5 dem_{it} + \in_{it} \quad (6)$$

where buildup=1 if Officer *i* is observed during FY07-FY09 and zero otherwise; M is a vector of two indicator variables representing combat MOS and another for non-combat MOS. Thus, β_3 represents the DiD estimate of the net change in job performance of USMC officers during the buildup period, as measured by the officer's FITREP average relative to his or her reporting senior's highest score. It is thus the before-after change in job performance between the treated groups (combat and non-combat) relative to the control group, on average and holding everything else constant.

The performance (perf) variable is a vector of independent variables including: ACT Score (act), Missing ACT Score (xact), GCT Score (gct), LSAT Score (lsat), Missing LSAT(xlsat), SAT Score (sat), Missing SAT (_sat), SAT interacted with Missing SAT (sat_sat), Class of Rifle Score (rifle_class=1 if first class, zero otherwise), AFQT Score (afqt), weight control status (weight_control=1 if an officer was ever assigned to weight control), PFT class (pft_class=1 if pft class is first class, zero otherwise, swim qual class (swim qual=1 if advanced, zero otherwise), binary variable for CD (cd=1 if accepted, zero otherwise), and missing swim class (xswim). The demographic (dem) variable consists of: Marital Status (marital_status), Number of Dependents (num_dep), and Sex (sex, male=1, zero otherwise).

6. Model 6: FY Average FITREPSs minus FY RS High Averages during the Drawdown FY10–FY13

Model 6 is a linear regression model estimating the effects of the drawdown on the FY average of the FITREP averages minus the FY RS high averages for a given MRO. Equation 7 represents the model.

$$avg\_hi_{it} = \beta_0 + \beta_1 drawdown_t + \beta_2 M_i + \beta_3 drawdown_t * M_i + \beta_4 perf_{it} + \beta_5 dem_{it} + \in_{it} \quad (7)$$

where drawdown=1 if Officer *i* is observed during FY10-13 and zero otherwise; M is a vector of two indicator variables representing combat MOS and another for non-combat MOS. Thus, β_3 represents the DiD estimate of the net change in promotion rates during the build-up period. It is the before-after change in CD between the treated groups (combat and non-combat) relative to the control group.

The performance (perf) variable is a vector of independent variables including: ACT Score (act), Missing ACT Score (xact), GCT Score (gct), LSAT Score (lsat), Missing LSAT(xlsat), SAT Score (sat), Missing SAT (_sat), SAT interacted with Missing SAT (sat_sat), Class of Rifle Score (rifle_class=1 if first class, zero otherwise), AFQT Score (afqt), weight control status (weight_control=1 if an officer was ever assigned to weight control), PFT class (pft_class=1 if pft class is first class, zero otherwise, swim qual class (swim qual=1 if advanced, zero otherwise), binary variable for CD (cd=1 if accepted, zero otherwise), and missing swim class (xswim). The demographic (dem) variable consists of: Marital Status (marital_status), Number of Dependents (num_dep), and Sex (sex, male=1, zero otherwise).

7. Model 7: Cancian and Klein Study Validation

Model 7 is a linear regression estimating the effects of the FY average of FITREPSs on GCT scores. The intent of this model is to re-examine Cancian and Klein's 2015 study determining USMC officer quality has declined due to a reduction in GCT exam scores over time. Cancian and Klein's study analyzes data from 1980 to 2014, while this research analyzes data from 2004 to 2015 only. Thus, the exercise here is not a full replication, but a re-examination of Cancian and Klein's findings with a focus on the quality of more recent and current officer populations. Model 7 can be seen in equation 8.

$$gct_{it} = \beta_0 + \beta_1 FY_t + \beta_2 avg\_this\_rpt_{it} + \beta_3 avg\_this\_rpt_{it} * FY_t + \beta_4 perf_{it} + \beta_5 dem_{it} + \in_{it} \quad (8)$$

where FY is a vector of indicator variables of all FYs represented in the dataset, FY04-FY15. β_2 corresponds to the association of GCT scores with the FY average of an MRO's FITREPSs. Thus, β_3 represents the year-to-year change in FITREP averages associated with year-to-year changes in GCT scores.

The performance (perf) variable is a vector of independent variables including: ACT Score (act), Missing ACT Score (xact), GCT Score (gct), LSAT Score (lsat), Missing LSAT(xlsat), SAT Score (sat), Missing SAT (_sat), SAT interacted with Missing SAT (sat_sat), Class of Rifle Score (rifle_class=1 if first class, zero otherwise), AFQT Score (afqt), weight control status (weight_control=1 if an officer was ever assigned to weight control), PFT class (pft_class=1 if pft class is first class, zero otherwise, swim qual class (swim qual=1 if advanced, zero otherwise), binary variable for CD (cd=1 if accepted, zero otherwise), MOS category for combat (combat=1 if combat MOS, zero otherwise), MOS category for non-combat (non-combat=1 if non-combat MOS, zero otherwise), and missing swim class (xswim). The demographic (dem) variable consists of: Marital Status (marital_status), Number of Dependents (num_dep), and Sex (sex, male=1, zero otherwise).

THIS PAGE INTENTIONALLY LEFT BLANK

IV. RESULTS

A. MODELS 1 AND 2: PROBIT CD DURING THE BUILDUP (FY07–FY09) AND DRAWDOWN (FY10–FY13)

1. Net Change in FY Average of FITREP Averages between the Buildup and Drawdown

The primary purpose of this thesis is to determine if there is a net change in the quality of USMC officers as a result of the transitional period of the buildup relative to the drawdown. One key variable is to view the change in the FY average of FITREPSs (avg_this_rpt) for a given officer relative to the probability of being selected for CD.

Table 1 displays the results of the effect of FY average of FITREPSs on CD selection for both periods. The full results, with all variables, are reported in Appendix A. Note that all other factors in Appendix A (e.g., test scores and demographics) are being held constant even as we compare across the treatment and control groups, in an effort to make an apples-to-apples comparison across the time periods. Estimates of Model 1 show that on average, 1 more point on the FY average FITREPSs during the buildup increases the probability of CD selection by 4.87 percentage points, for the treatment group relative to the control group in the same period. During the drawdown, Model 2 shows that on average, 1 more point on the FY average FITREPSs increases the probability of CD selection by 5.89 percentage points, again for the treatment relative to the control group in the same period. Both results are statistically significant at the 99 percent confidence level.

A general conclusion about the change in quality between both periods can be drawn by taking the difference between both models (triple difference). Table 1 shows that the effect of "avg_this_rpt" on CD selection during the drawdown minus the effect of "avg_this_rpt" during the buildup yields a 1.02 percentage point increase in the probability of being offered CD. In other words, the USMC has a higher reliance on FITREPs as a predictor of future success during the drawdown, relevant to the buildup.

Since different models are not guaranteed to yield the same results due to the observations not being identical, it is imperative to conduct a joint test between the two

models to determine if 1.02 percentage points are significantly different than zero. Simultaneously estimating Models 1 and 2 allows for this joint test. The test rejects the null hypothesis that there is no statistically significant difference in the effect of FITREP average scores on CD selection across the buildup versus drawdown (*p*-value of 0.0000); thus, the triple difference of 1.02 percentage points between the models is relevant and significantly different than 0.

Table 1. Models 1 and 2 FY Average FITREP Triple Difference

Labels	**Variable**	**Model 1: Probit Buildup FY07-13**	**Model 2: Probit Drawdown FY07-13**	
FY FITREP Average During Buildup (FY07-10)	avg_this_rpt build	0.0487		
FY FITREP Average During Drawdown (FY10-13)	avg_this_rpt draw		0.0589	
	Triple Difference			**0.0102**

The joint test verifies that the USMC has a greater focus on the FITREP during the drawdown for selecting officers to CD. This finding is consistent with the interpretation that with the greater demand for quantity (nearly 100 percent of officers were retained) during the buildup, the trade-off was a lower quality of officers. Meanwhile, during the drawdown period when only 55 percent were being retained, above average officers must have been selected to remain in service. To empirically establish whether the quality of the officer population did indeed change with end strength changes, the later set of Models examine differences in FITREPs as the dependent variable.

2. Net Change in FY Average of FITREPSs minus FY Average of RS Highs between the Buildup and Drawdown

An alternative measure of officer quality based on their job performance is the difference between the FY average of FITREPSs minus the FY average of reporting senior (RS) highs. The RS high is the highest FITREP written by an RS on any MRO of

the same rank. If the RS writes a report higher than the current high, the new highest FITREP is recorded for future reference. Thus, this measure of officer quality reflects a Marine officer's job performance relative to the best similarly-ranked officer the evaluator ever rated. Using this alternative measure of officer quality, I also test the hypothesis that the quality of USMC officers as a result of the transitional period of the buildup relative to the drawdown is significantly different from 0. The DiD for each model can be seen in the interaction variables with "avg_hi" and either the buildup or drawdown dummy variable.

Table 2 displays the results of the effect of the FY average FITREPs minus the FY average of RS highs on CD selection for both periods. The full results, with all variables, are reported in Appendix A. Note that similar to the previous results, all other factors in Appendix A (e.g., test scores and demographics) are being held constant even as we compare across the treatment and control groups, in an effort to make an apples-to-apples comparison across the time periods. Estimates of Model 1 show that on average, 1 more point on the FY average FITREP minus the FY average RS Highs during the buildup increases the probability of CD selection by 5.67 percentage points, for the treatment group relative to the control group in the same period. During the drawdown, Model 2 shows that on average, 1 more point on the FY average FITREPS minus the FY average of RS Highs increases the probability of CD selection by 9.12 percentage points, again for the treatment group relative to the control group in the same period. Both results are statistically significant at the 99 percent confidence level.

Table 2. Models 1 and 2 FY Average FITREPSs minus FY Average RS Highs Triple Difference

Label	**Variable**	**Model 1: Probit Buildup FY07-13**	**Model 2: Probit Drawdown FY07-13**	
FY FITREP Average minus FY Average RS High During Buildup (FY07-10)	avg_hi_build	0.0567		
FY FITREP Average minus FY Average RS High During Drawdown (FY10-13)	avg_hi_draw		0.0912	
	Triple Difference			**0.0345**

As with the previous estimates, a general conclusion about the change in quality between both periods can be drawn by taking the difference between both models (triple difference). Table 2 shows that the effect of “avg_hi” on CD selection during the drawdown minus the effect of “avg_hi” during the buildup yields a 3.45 percentage point increase in the probability of being offered CD. In other words, the USMCs reliance on FITREPs as a future predictor of success is significantly higher in the drawdown relative to the buildup. In this sense, CD was more selective based on FITREPs during the drawdown vs. the buildup.

Again, different models are not guaranteed to yield the same results due to the observations not being identical. It is imperative to conduct a joint test between the two models to determine if 3.45 percentage points are significantly different than zero. Simultaneously estimating Models 1 and 2 allows for this joint test. The test rejects the null hypothesis that there is no statistically significant difference in the effect of FITREP average scores minus average RS highs on CD selection across the buildup vs. drawdown (*p*-value of 0.0000); thus, the triple difference of 3.45 percentage points between the models are relevant and significantly different than 0. Note also that this triple difference estimator is three times larger in magnitude than the previous estimate.

3. Probability of CD for Combat MOS Group between the Buildup and Drawdown

Given the above results, it is useful to look at the change in probability of being offered CD amongst MOS groups. This can be conducted by taking a DiD in the model by including interaction variables for the MOS and period. As shown in Table 3, the combat MOS group during the buildup period were 5.54 percentage points more likely to be offered CD relative to the control group. Whereas, during the drawdown period the combat MOS group is 2.76 percentage points more likely to be accepted for CD than the control group. Both results are statistically significant at the 99 percent confidence level.

The triple difference for the combat MOS can then be formed by subtracting the results of Model 1 from Model 2. Table 3 shows the reduction in the probability of being offered CD in the drawdown is 2.78 percentage points relative to the buildup. Yet again,

a joint test must be conducted to determine the significance of this triple difference. The joint test yields a *p*-value of 0.0001, which rejects the null that the probability of being offered CD during both periods is the same. This is consistent with the earlier figures that nearly 100 percent of officers were retained during the buildup while selection of the combat MOS fell to 55 percent during the drawdown. The full results of the model can be seen in Appendix A.

Table 3. Probability of CD for Combat MOS Group between the Buildup and Drawdown

Label	Variable	Model 1: Probit Buildup FY07-13	Model 2: Probit Drawdown FY07-13	
Combat Interacted with Buildup (FY07-09)	comb_build	0.0554		
Combat Interacted with Drawdown (FY10-13)	comb_draw		0.0276	
	Triple Difference			**-0.0278**

4. Probability of CD for Non-Combat MOS Group between the Buildup and Drawdown

The results should also be analyzed for the non-combat MOS group. Once again, the difference-in-differences is estimated by including interaction variables for the MOS and period. As shown in Table 4, the non-combat MOS group during the buildup period is 1.93 percentage points more likely to be offered CD relative to the control group in the same period. Whereas, during the drawdown period the non-combat MOS group is 5.95 percentage points less likely to be accepted for CD than the control group. Both results are statistically significant at the 99 percent confidence level.

The triple difference for the non-combat MOS is estimated by subtracting the results of Model 1 from Model 2. The probability of being offered CD in the drawdown is lower by 7.88 percentage points for the non-combat group relative to the control group. Yet again, a joint test must be conducted to determine the significance of the triple difference. The joint test yields a *p*-value of 0.0000, which rejects the null hypothesis that

the probability of being offered CD during both periods is the same. Similar to the earlier figures for combat MOS, this is consistent with the fact that nearly 100 percent of officers were retained during the buildup and the non-combat MOS fell to 55 percent during the drawdown.

Table 4. Probability of CD for Non-Combat MOS Group between the Buildup and Drawdown

Label	**Variable**	**Model 1: Probit Buildup FY07-13**	**Model 2: Probit Drawdown FY07-13**	
Non-Combat Interacted with Buildup (FY07-09)	non-comb_build	0.0193		
Non-Combat Interacted with Drawdown (FY10-13)	non-comb_draw		-0.0595	
	Triple Difference			**-0.0788**

B. MODELS 3 AND 4: FY AVERAGE FITREPSS DURING THE BUILDUP (FY07-FY09) VERSUS THE DRAWDOWN (FY10–FY13)

The above results establish that the USMC put a statistically greater emphasis on the FITREP during the drawdown for selecting officers to CD. This finding is consistent with the interpretation that with the greater demand for quantity during the buildup, the trade-off was a lower quality of officers. To empirically determine the change in officer quality, Models 3 and 4 estimate the net change in the FY average FITREPSs between the buildup and drawdown periods by MOS groups.

1. Combat MOS Group Net Change in FY Average of FITREPSs between the Periods

Table 5 displays the results of the change in FY average FITREPS experienced by the combat MOS group during the buildup and drawdown periods. The full results, with all variables, are reported in Appendix B. Note that all other factors in Appendix B (e.g., test scores and demographics) are being held constant even as we compare across the

treatment and control groups, in an effort to make an apples-to-apples comparison across the time periods.

Model 3 estimates the combat MOS group received FITREPS that, on average, were 0.4592 points higher than the control group in the same period. Recall that the average FITREP scores are on a 1 to 7 point scale, and 0.4592 is about a quarter of a standard deviation. During the drawdown, the combat group, on average receives FITREPS that are 0.4913 points higher, relative to the control group in the same period. Again, 0.4913 points is about a quarter of a standard deviation. Both results are statistically significant at the 99 percent confidence level.

As before, by taking the difference between both models (triple difference), I can form estimates of the change in quality between both periods. Table 5 shows the effect of having a combat MOS on the FY average FITREP during the drawdown minus the effect of having a combat MOS on the FY average FITREP during the buildup yields a 0.0321 point increase in the average FITREP score. In other words, the quality of USMC officers as measured by the positive change in average FITREP scores is higher in the drawdown relative to the buildup.

Again, different models are not guaranteed to yield the same results due to the observations not being identical. Simultaneously estimating Models 3 and 4 allows for a joint test between the two models to determine if 0.0321 points are significantly different than zero. The test rejects the null hypothesis that there is no statistically significant difference in the quality of officers with a combat MOS across the buildup vs. drawdown (p-value of 0.0223); thus the triple difference of 0.0321 points between the models are relevant and significantly different than 0. This is consistent with the primary hypothesis that quality changed between the periods and the drawdown period possesses higher quality combat MOS officers, on average.

Table 5. Combat MOS Group Net Change in FY Average of FITREPSs between the Periods

Label		**Model 3: Average This Report Buildup FY07-13**	**Model 4: Average This Report Drawdown FY07-13**	
Combat Interacted with Buildup (FY07-09)	combat build	0.4592		
Combat Interacted with Drawdown (FY10-13)	combat draw		0.4913	
	Triple Difference			**0.0321**

2. Non-Combat MOS Group Net Change in FY Average of FITREPSs between the Periods

Table 6 displays the results of the change in FY average FITREPS experienced by the non-combat MOS group during the buildup and drawdown periods. The full results, with all variables, are reported in Appendix B. Note that all other factors in Appendix B (e.g., test scores and demographics) are being held constant even as we compare across the treatment and control groups, as with the previous models.

Model 3 estimates the non-combat MOS group during the buildup received FITREPS that, on average were 0.381 points higher, relative to the control group in the same period. Recall that the average FITREP scores are on a 1 to 7 point scale, and 0.381 is also about a quarter of a standard deviation. During the drawdown, the non-combat group, on average receives FITREPS that are 0.4644 points higher, relative to the control group in the same period. Again, 0.4644 points is about a quarter of a standard deviation. Both results are statistically significant at the 99 percent confidence level.

To estimate the change in quality between both periods, I again take the difference between both models (triple difference). Table 6 shows that the effect of having a non-combat MOS on the FY average FITREP during the drawdown minus the effect of having a non-combat MOS on the FY average FITREP during the buildup yields a 0.0834 point increase in the average FITREP score. In other words, the quality of

USMC officers as measured by the positive change in average FITREP scores is higher in the drawdown relative to the buildup for the non-combat MOS group.

As before, I conduct a joint test to see if 0.0834 points are significantly different from zero. The test rejects the null hypothesis that there is no statistically significant difference in the quality of officers with a non-combat MOS across the buildup vs. drawdown (*p*-value of 0.0000); thus the triple difference of 0.0834 points between the Models 3 and 4 are relevant and significantly different than 0. This is consistent with the primary hypothesis that quality changed between the periods and the drawdown period is characterized by higher quality non-combat MOS officers, on average.

Table 6. Non-Combat MOS Group Net Change in FY Average of FITREPSs between the Periods

Label		**Model 3: Average This Report Buildup FY07-13**	**Model 4: Average This Report Drawdown FY07-13**	
Non-Combat Interacted with Buildup (FY07-09)	non-comb build	0.381		
Non-Combat Interacted with Drawdown (FY10-13)	non-comb draw		0.4644	
	Triple Difference			**0.0834**

C. MODELS 5 AND 6: REGRESSION OF FY AVERAGE FITREPSS MINUS FY AVERAGE REPORT HIGHS DURING THE BUILDUP (FY07–FY09) VERSUS THE DRAWDOWN (FY10–FY13)

Models 5 and 6 analyze the net change in the FY average of FITREPSs minus the FY average of the RS high, “avg_hi,” relative to each MOS group over the duration of the buildup and drawdown periods. The “avg_hi” should be a negative number unless, in the unlikely event, that all FITREPSs in the population were written above the RS high.

1. Combat MOS Group Net Change in FY Average of FITREPSs minus the FY Average of RS High between the Two Periods

Table 7 displays the results of the change in FY average FITREPS minus FY average RS highs (avg_hi) experienced by the combat MOS group during the buildup and drawdown periods. The full results, with all variables, are reported in Appendix C. Note that all other factors in Appendix C (e.g., test scores and demographics) are also being held constant even as we compare across the treatment and control groups. Model 5 estimates results of “avg_hi” for the combat MOS group were 0.0491 points lower, relative to the control group in the same period. During the drawdown, the combat group, on average experiences a 0.0403 reduction in “avg_hi,” relative to the control group in the same period.

To draw a general conclusion about the change in quality between both periods, I take the difference between both models (triple difference). Table 7 shows 0.0088 “avg_hi” point increase in the difference of having a combat MOS during the drawdown vs. buildup. In other words, the quality of USMC officers as measured by the positive change in “avg_hi” is higher in the drawdown relative to the buildup. Interpreting the results, a positive number implies the “avg_hi” becomes closer to zero. For the difference in the models to become less negative, job performance ratings are such that they are closer to or exceeding the RS’s high score at the time of processing, signifying higher officer quality during the drawdown compared to the buildup.

Again, different models are not guaranteed to yield the same results due to the observations not being identical. It is imperative to conduct a joint test between the two models to determine if 0.0088 points are significantly different than zero. Simultaneously estimating Models 5 and 6 allows for this joint test. The test fails to reject the null hypothesis that there is no statistically significant difference in the quality of officers with a combat MOS across the buildup vs. drawdown (*p*-value of 0.1998); thus quality using this measure is statistically identical during this period.

The results of this joint test is not consistent with the primary hypothesis. However, this could also be due to the fact that combat RSs did not change their high and/or average FITREP scoring across the periods and thus internally prevented inflation.

Again, the assumption here is that FITREPS are the best measure of quality and are not inflated across periods. As discussed previously, the RV is an alternative measure that takes into account inflation occurring due to different RS reporting styles over time. This research did not use RVs as measures of quality, but does recommend it as follow on research to further validate the primary hypothesis. Nonetheless, the preponderance of the estimates here are consistent with the primary hypothesis.

Table 7. Combat MOS Group Net Change in FY Average of FITREPSs minus the FY Average of RS High between the Two Periods

Label	**Variable**	**Model 5: Report Average Buildup FY07-13**	**Model 6: Report Average Drawdown FY07-13**	
Combat Interacted with Buildup (FY07-09)	comb_build	-0.0491		
Combat Interacted with Drawdown (FY10-13)	comb_draw		-0.0403	
	Triple Difference			**0.0088**

2. Non-Combat MOS Group Net Change in FY Average of FITREPSs minus the FY Average of RS High between the Two Periods

Table 8 displays the results of the change in FY average FITREPS minus FY average RS highs (avg_hi) experienced by the non-combat MOS group during the buildup and drawdown periods. The full results, with all variables, are reported in Appendix C. As with the previous estimates, all other factors in Appendix C (e.g., test scores and demographics) are being held constant even as we compare across the treatment and control groups. Model 5 estimates results of “avg_hi” for the non-combat MOS group were 0.0323 points lower, relative to the control group in the same period. During the drawdown, the non-combat group, on average experiences a 0.0173 reduction in “avg_hi,” relative to the control group in the same period.

A general conclusion about the change in quality between both periods can be drawn by taking the difference between both models (triple difference). Table 8 shows the effect of having a non-combat MOS on the "avg_hi" during the drawdown minus the buildup yields a 0.015 point increase. In other words, the quality of USMC officers as measured by the positive change in "avg_hi" is higher in the drawdown relative to the buildup.

Again, different models are not guaranteed to yield the same results due to the observations not being identical. Simultaneously estimating Models 5 and 6 allows for a joint test between the two models to determine if 0.015 points are significantly different than zero. The test rejects the null hypothesis that there is no statistically significant difference in the quality of officers with a non-combat MOS across the buildup vs. drawdown (*p*-value of 0.015). Thus, the triple difference of 0.015 points between the models is relevant and significantly different than 0. This is consistent with the primary hypothesis that quality changed between the periods and the drawdown period had higher quality non-combat MOS officers.

Table 8. Non-Combat MOS Group Net Change in FY Average of FITREPSs minus the FY Average of RS Highs between the two Periods

Label	**Variable**	**Model 5: Report Average Buildup FY07-13**	**Model 6: Report Average Drawdown FY07-13**	
Combat Interacted with Buildup (FY07-09)	non_comb_build	-0.0323		
Combat Interacted with Drawdown (FY10-13)	non_comb_draw		-0.0173	
	Triple Difference			**0.015**

D. MODEL 7: CANCIAN AND KLEIN 2015 STUDY VALIDATION

Model 7 is a robustness check of the Cancian and Klein 2015 study which argued that USMC officer quality has declined due to a decline in GCT scores. The pertinent results can be seen in Table 10, while comprehensive results can be seen Appendix D. Note, too, that all other factors in Appendix D are being held constant. The variables "_Ify_2005" to "_Ify_2015" represent the percent change in GCT score on average in each FY relative to 2004. The results show the scores were declining throughout the period, thus verifying Cancian and Klein's finding that GCT scores had declined. Cancian and Klein argue that a reduction in GCT scores corresponds to a reduction in officer quality; thus it is important to verify the correlation of GCT scores with output based measures of quality, like the FITREP scores used in this research. As shown in Table 9, there is a very weak correlation between GCT and "avg_this_rpt" or "avg_hi." While positive, both measures of officer quality have a correlation of less than 0.05 with GCT scores. This weak correlation also indicates that GCT scores are not predictive of officer productivity and job performance.

Table 9. GCT Correlation Matrix

	gct	**avg_this_rpt**	**avg_hi**	**sat**
gct	1.0000			
avg_this_rpt	0.0189	1.0000		
avg_hi	0.0366	-0.3429	1.0000	
sat	0.0750	0.0666	0.0128	1.0000

Cancian and Klein bases their report on the assumption that GCT is a significant measure of officer quality. I argue instead that the best measure of quality is the FITREP, since it is the most comprehensive, output-based, and job performance evaluation tool for an officer. As seen in Table 9, there is no strong correlation between GCT and qualitative variables; therefore, making it difficult to utilize GCT as a measure of quality. This research used the FY average FITREPSs for an MRO and interacted it with the FY (avg_this_rpt * FY). Using the DiD method, there is not a single year where the FY average FITREPSs declined due to a decline in GCT scores. Therefore, Cancian and

Klein's 2015 study is partially correct that GCT scores declined over time. However, GCT scores do not comprehensively define the quality of an officer. Cancian and Klein did not have the pertinent data to make a definitive statement that officer quality had declined.

Table 10. Cancian and Klein 2015 Model Results

VARIABLES	LABELS	Model 7
_Ify_2005	fy==2005	0.4002
		[0.3130]
_Ify_2006	fy==2006	-0.8782***
		[0.2915]
_Ify_2007	fy==2007	-0.9009***
		[0.2901]
_Ify_2008	fy==2008	-2.1997***
		[0.2932]
_Ify_2009	fy==2009	-2.2724***
		[0.2930]
_Ify_2010	fy==2010	-2.4481***
		[0.8846]
_Ify_2011	fy==2011	-2.7409***
		[0.8843]
_Ify_2012	fy==2012	-2.8174***
		[0.8842]
_Ify_2013	fy==2013	-3.1527***
		[0.8852]
_Ify_2014	fy==2014	-3.4220***
		[0.8855]
_Ify_2015	fy==2015	-3.4896***
		[0.8826]
avg_this_rpt	FY avg of rpt avg	0.1661***
		[0.0473]
_IfyXavg_2005	(fy==2005)*avg_this_rpt	-0.0064
		[0.0681]
_IfyXavg_2006	(fy==2006)*avg_this_rpt	-0.0563
		[0.0623]
_IfyXavg_2007	(fy==2007)*avg_this_rpt	-0.0946
		[0.0621]
_IfyXavg_2008	(fy==2008)*avg_this_rpt	-0.0913
		[0.0621]
_IfyXavg_2009	(fy==2009)*avg_this_rpt	-0.1199*
		[0.0619]
_IfyXavg_2010	(fy==2010)*avg_this_rpt	-0.0923
		[0.0615]
_IfyXavg_2011	(fy==2011)*avg_this_rpt	-0.0648

VARIABLES	LABELS	Model 7
		[0.0610]
_IfyXavg_2012	(fy==2012)*avg_this_rpt	-0.1227**
		[0.0610]
_IfyXavg_2013	(fy==2013)*avg_this_rpt	-0.1045*
		[0.0614]
_IfyXavg_2014	(fy==2014)*avg_this_rpt	-0.0968
		[0.0615]
_IfyXavg_2015	(fy==2015)*avg_this_rpt	-0.1337**
		[0.0596]

E. ROBUSTNESS CHECKS

The models in the previous sections were validated by several robustness checks. The data collected from MMRP-30 and TFDW spanned FY04–FY15. The actual buildup and drawdown periods were mostly in FY07-FY13. This can be seen in the change in officer population in Figure 3. Initially the models included all years FY04–FY15, then FY07–FY15, and FY07–FY13. The results for all models were consistent across these various samples. Given the robustness of these results, I reported the models using only FY07–FY13 as it represents the actual years of the buildup and drawdown to test the primary hypothesis. While the precise dates of the buildup and drawdown could include adjacent years, FY07–FY13 is representative of the actual periods and the results were similar throughout all FYs in the data.

As additional robustness checks, the models were estimated using alternative control groups. The assumption in the main results is that pilots and lawyers make the best control group, since their CD selection rates are at or near 100 percent throughout the data. To validate this assumption, I created three control groups for testing. As in the models above, a DiD model for the buildup and drawdown periods were estimated using the alternative control groups. Subsequently, a triple difference was calculated between the models to determine the net change of the results between the periods. This allowed me to determine if each of the three control groups had similar results.

First, a control group with pilots and lawyers (control) only was created; these are the results reported above. Second, a control group with lawyers only (control_lawyer)

was created to determine if the results were consistent with the first. Using the triple difference estimate, the results of the lawyers only control group were similar and consistent with the first control group. Third, a control group with pilots (control_pilots) only was created. Interestingly, the triple difference estimate with pilots only as a control group were inconsistent with the first two and the sign of the coefficients were opposite. The results showed the USMC became more selective of pilots during the drawdown than combat MOS, relative to the buildup. Meanwhile, there was no significant triple difference between pilots and non-combat across the time periods. While these results suggest lawyers are mostly driving the triple difference results given previously, I still felt the control group of pilots and lawyers were consistent and representative of the historical CD selection results and chose to keep it as my control group for the main part of the analysis.

V. CONCLUSION, RECOMMENDATIONS, AND FUTURE RESEARCH

A. SUMMARY AND CONCLUSION

The purpose of this research is to investigate whether there is a change in the quality of USMC officers as a result of the buildup and the subsequent drawdown of forces. This was accomplished by using multiple probit and multivariate regression models to analyze changes in CD selection and average FITREPs over time. The results of this research may be used as a foundation for future conflicts where drastic end strength changes may be warranted. It highlights that better methods for shaping the force should be researched and implemented due to the trade-off between quantity and quality of officers during the drawdown period evaluated.

I argue that FITREPs are the best available measure of quality for an officer, as it is an output-based measure of job performance rather than a measure of input such as AFQT or other raw ability measure. First, I find conclusive evidence that CD was more selective in the drawdown than the buildup. This should be obvious in the raw unconditional figures, since nearly 100 percent of officers were retained in the buildup and only 55 percent in the drawdown. Second, I find systematic evidence of an increase in the average FITREP score between the periods, even when I examine within USMC officers using a treatment-control research design. These findings further solidify the first finding that quality increased during the drawdown, although not statistically significant for combat MOSs, but still economically relevant. Additionally, the difference between the average FITREP score and RS high narrowed between the periods. This signifies more FITREPs were written at or above the RS high, further indicating higher officer quality in the drawdown.

Third, I check the validity of Cancian and Klein's 2015 research which claimed that Marine officer quality has declined as evidenced by a reduction in GCT scores. Fortunately, their study became mainstream during the writing of this thesis, thus allowing me to use my data to test their results. The results of my own analysis were conclusive that GCT scores had declined over time; however, officer FITREPs did not

and other measures of officer ability such as SAT, ACT, AFQT scores did not. Part of the issue is that different officer populations over time are taking and reporting these tests, making these input-based measures a complicated metric for evaluating officer quality. In fact, using a multivariate regression model with DiD estimators, overall officer quality never had a reduction from FY04-FY15 as measured by GCT scores, and there is a weak correlation between GCT scores and FITREP scores. I interpret this as evidence that Cancian and Klein's 2015 study is inconsistent with any changes in the actual quality of USMC officers for the years 2004 to 2015.

Buildups and drawdowns of the U.S. military force are necessary for maintaining national security. However, military leaders are constrained by presidential and congressional end strength decisions and are required to act expeditiously to meet end strength mandates by the end of each FY. Even as military leaders serve at the president's will and strive to achieve his decision in the desired manner, it is apparent that they must make trade-off decisions between quantity and quality to meet political end strength requirements.

Johnston (2015) discusses the trade-off that the USMC must make between fully staffing an infantry unit with enough time to train and become cohesive prior to a deployment vs. not (Johnston, 2015, p. 87). This research has identified and quantified the trade-off USMC makes between quality and quantity of USMC Officers, with quantity being the focus during the buildup, which in turn forces very tight qualitative standards during the drawdown. It takes multiple periods to fully mitigate the effects of these policies and return the system to a steady state of like-quality officers.

While the system churns toward a steady state, a significant share of quality officers are not retained, but low-quality officers promoted during the buildup are still in the system and difficult to remove. Bacolod (2007), analyzes the decline in teacher quality due to expanded access to professional jobs for women. Bacolod states, "Thus, changes in average relative wages [in teaching vs professional jobs] affect not only the decision to enter teaching but also the average quality of teachers" (Bacolod, 2007, p. 745). The theoretical model presented in Bacolod (2007) has particular implications for USMC officer quality. During the drawdown, high-quality officers with attractive

alternative labor market opportunities in the civilian sector are the ones most likely to accept voluntary separations compared to lower quality officers.

The Marine Corps rarely incentivizes high-quality officers with early promotions that may deter their decision to accept a more lucrative, non-military job. Thus, it may become even more difficult to achieve a steady state of officer quality due to secondary and tertiary effects of the buildup and drawdown retention policies. In conclusion, a one-size-fits-all force structuring policy may not create the best quality of officer force, but a flexible, precise approach similar to civilian sector policies, may improve the quality of officers. Without changes to its current manpower policies, Marine Corps officers may mirror the decline experienced in teacher quality.

B. RECOMMENDATIONS

I recommend that past practices utilized to grow and reduce the military force be examined with particular attention paid to the quality trade-off. The review should ensure plans to change the size of the force are done gradually, thus reducing the need for drastic retention, assessment, or retirement methods needed on each end of the end strength spectrum. Competitive CD may be kept to ensure the lowest quality officers are being removed at all times and across all services. Alternatively, competitive CD could be removed if the Title 10 requirement for captain promotions were made competitive. For instance O-4 and O-5 promotions are 80 percent and 70 percent respectively. A +/- 10 percent is normally allowed for the services to shape the officer force as needed. Creating a 90 percent +/- 10 percent requirement for captains would provide force-shaping methods and thwart the need for competitive CD. This would prevent an excess of non-promotable officers at any rank, reduce the need for drastic force-changing measures, and allow the force to be shaped over time. Obviously any changes would have to be palatable to congress, since Title 10 would have to be modified.

C. FUTURE RESEARCH

1. Correct Population Targeted in Early/Voluntary Separation Programs

This research is concerned with the change of officer quality during drastic end strength increases and decreases. If the buildup retains almost all officers and the drawdown retains the top 55 percent, then the buildup officers of all quality remain in the system during the drawdown. They have likely been promoted into the field grade ranks. Incentives must be used to entice a certain target populations to voluntarily leave the service.

It is generally assumed that lower quality officers would likely not accept early or voluntary separation packages due to a level of safety or comfort in their current position and lack of alternative labor market opportunities. In contrast, high-quality officers would be more likely to accept these offers and enter the civilian sector. A study should be conducted to analyze the effects of early or voluntary retirement options on the quality of the USMC officer corps. This could be conducted with the FITREP data used in this study and analyze the net change in quality based on separations codes. This will ensure the correct population is being targeted and reduce the loss of quality officers.

2. RS versus RO Remarks for Officer Advancement

This research focused solely on the RS evaluations of an MRO. The RS is directly subordinate to the RO and the RO is required to evaluate an MRO and either concur or not concur with the RS's evaluation. It is generally accepted that RO comments are more valuable for career progression, since a RO is likely to have evaluated more officers throughout their career. A study should be conducted to determine whether the remarks of the RS and RO are of equal importance for the career progression of an officer. This could be conducted by using the FITREP data in this research and evaluate RO and RS averages at different milestones in an MRO's career. The various milestones may be CD, promotions, special school selections, and so forth.

3. Quality across Race/Ethnicity

In recent years there has been an initiative to recruit diversity throughout the military. This is a very noble cause that seeks to create a culturally diverse professional military. The military needs quality, versatile members to fill its ranks. Recruiting diversity standards that ignore standards of quality may create a less-capable force. Research should be conducted to determine quality disparities between race and ethnicity to help determine if the current recruiting constraints are beneficial or harmful. This could be conducted with FITREP, race, and ethnicity data.

4. Relative Value (RV)

RVs are generated to help prevent inflation by RSs and ROs. The RS or RO must evaluate at least three officers of the same rank to gain a RV. A cumulative RV is also generated to track the officer's quality over time relative to their peers. The researcher must not forget about the recency effect and the fact that a RS is likely to write a current MRO higher than an MRO of past. This research did not use RV due to a lot of missing data and the process of creating a FY RV average was not very clean. The FITREP data could be used to conduct additional research that would serve as a robustness check for this thesis.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. PROBIT MODELS 1 AND 2 COMPARATIVE EFFECTS OF THE BUILDUP AND DRAWDOWN ON THE PROBABILITY OF BEING SELECTED FOR CAREER DESIGNATION

VARIABLES	LABELS	Model 1: Probit Buildup FY07-FY13	Model 2: Probit Drawdown FY07-FY13
fy	Fiscal year	-0.0044***	-0.0044***
		[0.0013]	[0.0013]
avg_this_rpt	FY avg of rpt avg	0.0588***	0.0488***
		[0.0009]	[0.0011]
_IbuiXavg_t_1		-0.0101***	
		[0.0015]	
avg_hi	FY rpt avg minus rs high	0.0912***	0.0567***
		[0.0043]	[0.0047]
_IbuiXavg_h_1		-0.0345***	
		[0.0063]	
combat	Combat MOS Group	0.0306***	0.0542***
		[0.0042]	[0.0048]
combat_buildup	combat*buildup	0.0248***	
		[0.0064]	
non_combat	Non-Combat MOS Group	-0.0544***	0.0279***
		[0.0039]	[0.0045]
non_combat_buildup	non_combat*buildup	0.0737***	
		[0.0046]	
buildup	FY05-FY09	-0.0904***	
		[0.0079]	
marital_status	Married Binary	0.0441***	0.0441***
		[0.0037]	[0.0037]
num_dep	Number of Dependents	0.0716***	0.0716***
		[0.0012]	[0.0012]
sex	Male=1	-0.0442***	-0.0442***
		[0.0043]	[0.0043]
act	ACT Score	0.0001***	0.0001***
		[0.0000]	[0.0000]
xact	Missing ACT	0.2107***	0.2107***
		[0.0104]	[0.0104]
gct	GCT Score	0.0024***	0.0024***
		[0.0001]	[0.0001]
lsat	LSAT Score	0.0001**	0.0001**
		[0.0000]	[0.0000]
xlsat	Missing LSAT	0.1395***	0.1395***
		[0.0178]	[0.0178]
sat	SAT Score	-0.0001***	-0.0001***
		[0.0000]	[0.0000]

VARIABLES	LABELS	Model 1: Probit Buildup FY07-FY13	Model 2: Probit Drawdown FY07-FY13
_sat	Missing SAT	-0.1781***	-0.1781***
		[0.0016]	[0.0016]
sat_sat	sat * _sat	0.0004***	0.0004***
		[0.0000]	[0.0000]
rifle_class	Expert=1, 0 Otherwise	0.0369***	0.0369***
		[0.0028]	[0.0028]
afqt	AFQT_SCORE	-0.0008***	-0.0008***
		[0.0000]	[0.0000]
weight_control	1=ever assigned to weight control	0.0679***	0.0679***
		[0.0119]	[0.0119]
pft_class	1st Class=1, 0 Otherwise	0.0664***	0.0664***
		[0.0037]	[0.0037]
swim	Advanced =1, 0 Otherwise	-0.0255***	-0.0255***
		[0.0033]	[0.0033]
xswim	Missing Swim Class	-0.1720***	-0.1720***
		[0.0198]	[0.0198]
_IdraXavg_t_1	drawdown*avg_this_rpt		0.0101***
			[0.0015]
_IdraXavg_h_1	drawdown*avg minus RS high		0.0345***
			[0.0063]
combat_drawdown	combat*drawdown		-0.0266***
			[0.0074]
non_combat_drawdown	non_combat*drawdown		-0.0874***
			[0.0065]
drawdown	FY10-FY15		0.0904***
			[0.0079]
Observations		107,034	107,034
Standard errors in brackets			
*** p<0.01, ** p<0.05, * p<0.1			

APPENDIX B. REGRESSION MODELS 3 AND 4 ANALYSIS OF THE NET CHANGE IN FY REPORT AVERAGES BETWEEN THE BUILDUP AND DRAWDOWN

VARIABLES	LABELS	Model 3: Average This Report Buildup FY07-FY13	Model 4: Average This Report Drawdown FY07-FY13
fy	Fiscal year	0.0126**	0.0126**
		[0.0058]	[0.0058]
_Icombat_1	combat==1	0.4913***	0.4592***
		[0.0197]	[0.0239]
_Ibuildup_1		-0.9526***	
		[0.2075]	
_IcomXbui_1_1		-0.0321	
		[0.0306]	
_Inon_comba_1	non_combat==1	0.4644***	0.3810***
		[0.0174]	[0.0208]
o._Ibuildup_1		-	
_InonXbui_1_1		-0.0834***	
		[0.0262]	
cd	CD binary	1.0871***	1.0871***
		[0.0145]	[0.0145]
marital_status	Married Binary	0.0745***	0.0745***
		[0.0166]	[0.0166]
num_dep	Number of Dependents	0.0471***	0.0471***
		[0.0049]	[0.0049]
sex	Male=1	-0.1173***	-0.1173***
		[0.0250]	[0.0250]
act	ACT Score	0.0003	0.0003
		[0.0002]	[0.0002]
xact	Missing ACT	0.6708***	0.6708***
		[0.0379]	[0.0379]
lsat	LSAT Score	-0.0004***	-0.0004***
		[0.0002]	[0.0002]
xlsat	Missing LSAT	-0.4389***	-0.4389***
		[0.0688]	[0.0688]
sat	SAT Score	0.0002	0.0002
		[0.0002]	[0.0002]
_sat	Missing SAT	0.0197	0.0197
		[0.1140]	[0.1140]
sat_sat	sat * _sat	0.0001	0.0001
		[0.0002]	[0.0002]
rifle_class	Expert=1, 0 Otherwise	-0.0615***	-0.0615***
		[0.0126]	[0.0126]

VARIABLES	LABELS	Model 3: Average This Report Buildup FY07-FY13	Model 4: Average This Report Drawdown FY07-FY13
afqt	AFQT_SCORE	-0.0019***	-0.0019***
		[0.0002]	[0.0002]
dlab	DLAB Score	0.0002	0.0002
		[0.0001]	[0.0001]
o.xdlab	Missing DLAB = o,	-	-
pft_class	1st Class=1, 0 Otherwise	-0.0982***	-0.0982***
		[0.0198]	[0.0198]
xpft_class	Missing PFT Class	-1.0675***	-1.0675***
		[0.2071]	[0.2071]
swim	Advanced =1, 0 Otherwise	-0.0592***	-0.0592***
		[0.0151]	[0.0151]
xswim	Missing Swim Class	-0.2799***	-0.2799***
		[0.0724]	[0.0724]
weight_control	1=ever assigned to weight control	-0.3453***	-0.3453***
		[0.0709]	[0.0709]
gct	GCT Score	0.0029***	0.0029***
		[0.0006]	[0.0006]
_Idrawdown_1	drawdown==1		0.9526***
			[0.2075]
_IcomXdra_1_1	combat*drawdown		0.0321
			[0.0306]
o._Idrawdown_1	drawdown==1 = o,		-
_InonXdra_1_1	non_combat*drawdown		0.0834***
			[0.0262]
Constant	Constant	-23.4517**	-24.4043**
		[11.7561]	[11.7348]
Observations		107,034	107,034
R-squared		0.092	0.092

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

APPENDIX C. MODELS 5 AND 6 FY AVERAGE OF FITREPSS MINUS FY AVERAGE OF RS HIGH REPORTS BETWEEN THE BUILDUP AND DRAWDOWN

VARIABLES	LABELS	Model 5: Report Average Buildup FY07-FY13	Model 6: Report Average Drawdown FY07-FY13
fy	Fiscal year	0.0110***	0.0110***
		[0.0013]	[0.0013]
_Icombat_1	combat==1	-0.0403***	-0.0491***
		[0.0043]	[0.0053]
_Ibuildup_1		0.2208***	
		[0.0457]	
_IcomXbui_1_1		-0.0088	
		[0.0067]	
_Inon_comba_1	non_combat==1	-0.0173***	-0.0323***
		[0.0038]	[0.0046]
o._Ibuildup_1		-	
_InonXbui_1_1		-0.0150***	
		[0.0058]	
marital_status	Married Binary	0.0262***	0.0262***
		[0.0037]	[0.0037]
num_dep	Number of Dependents	0.0132***	0.0132***
		[0.0011]	[0.0011]
sex	Male=1	-0.0144***	-0.0144***
		[0.0055]	[0.0055]
cd	CD binary	-0.0362***	-0.0362***
		[0.0032]	[0.0032]
act	ACT Score	-0.0000	-0.0000
		[0.0000]	[0.0000]
xact	Missing ACT	-0.0448***	-0.0448***
		[0.0083]	[0.0083]
lsat	LSAT Score	-0.0001**	-0.0001**
		[0.0000]	[0.0000]
xlsat	Missing LSAT	0.0312**	0.0312**
		[0.0152]	[0.0152]
sat	SAT Score	0.0000	0.0000
		[0.0000]	[0.0000]
_sat	Missing SAT	-0.1239***	-0.1239***
		[0.0251]	[0.0251]
sat_sat	sat * _sat	0.0001	0.0001
		[0.0000]	[0.0000]
rifle_class	Expert=1, 0 Otherwise	0.0188***	0.0188***
		[0.0028]	[0.0028]

VARIABLES	LABELS	Model 5: Report Average Buildup FY07-FY13	Model 6: Report Average Drawdown FY07-FY13
afqt	AFQT_SCORE	-0.0002***	-0.0002***
		[0.0000]	[0.0000]
dlab	DLAB Score	0.0003***	0.0003***
		[0.0000]	[0.0000]
o.xdlab	Missing DLAB = o,	-	-
pft_class	1st Class=1, 0 Otherwise	0.0717***	0.0717***
		[0.0044]	[0.0044]
xpft_class	Missing PFT Class	0.2678***	0.2678***
		[0.0456]	[0.0456]
swim	Advanced =1, 0 Otherwise	-0.0085**	-0.0085**
		[0.0033]	[0.0033]
xswim	Missing Swim Class	0.0142	0.0142
		[0.0160]	[0.0160]
weight_control	1=ever assigned to weight control	-0.0984***	-0.0984***
		[0.0156]	[0.0156]
gct	GCT Score	0.0016***	0.0016***
		[0.0001]	[0.0001]
_Idrawdown_1	drawdown==1		-0.2208***
			[0.0457]
_IcomXdra_1_1	combat*drawdown		0.0088
			[0.0067]
o._Idrawdown_1	drawdown==1 = o,		-
_InonXdra_1_1	non_combat*drawdown		0.0150***
			[0.0058]
Constant	Constant	-22.9117***	-22.6909***
		[2.5902]	[2.5855]
Observations		107,034	107,034
R-squared		0.021	0.021

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

APPENDIX D. MODEL 7: CANCIAN AND KLEIN 2015 ROBUSTNESS CHECK

VARIABLES	LABELS	Model 7
_Ify_2005	fy==2005	0.4002
		[0.3130]
_Ify_2006	fy==2006	-0.8782***
		[0.2915]
_Ify_2007	fy==2007	-0.9009***
		[0.2901]
_Ify_2008	fy==2008	-2.1997***
		[0.2932]
_Ify_2009	fy==2009	-2.2724***
		[0.2930]
_Ify_2010	fy==2010	-2.4481***
		[0.8846]
_Ify_2011	fy==2011	-2.7409***
		[0.8843]
_Ify_2012	fy==2012	-2.8174***
		[0.8842]
_Ify_2013	fy==2013	-3.1527***
		[0.8852]
_Ify_2014	fy==2014	-3.4220***
		[0.8855]
_Ify_2015	fy==2015	-3.4896***
		[0.8826]
avg_this_rpt	FY avg of rpt avg	0.1661***
		[0.0473]
_IfyXavg_2005	(fy==2005)*avg_this_rpt	-0.0064
		[0.0681]
_IfyXavg_2006	(fy==2006)*avg_this_rpt	-0.0563
		[0.0623]
_IfyXavg_2007	(fy==2007)*avg_this_rpt	-0.0946
		[0.0621]
_IfyXavg_2008	(fy==2008)*avg_this_rpt	-0.0913
		[0.0621]
_IfyXavg_2009	(fy==2009)*avg_this_rpt	-0.1199*
		[0.0619]
_IfyXavg_2010	(fy==2010)*avg_this_rpt	-0.0923
		[0.0615]
_IfyXavg_2011	(fy==2011)*avg_this_rpt	-0.0648
		[0.0610]
_IfyXavg_2012	(fy==2012)*avg_this_rpt	-0.1227**
		[0.0610]
_IfyXavg_2013	(fy==2013)*avg_this_rpt	-0.1045*
		[0.0614]

VARIABLES	LABELS	Model 7
_IfyXavg_2014	(fy==2014)*avg_this_rpt	-0.0968
		[0.0615]
_IfyXavg_2015	(fy==2015)*avg_this_rpt	-0.1337**
		[0.0596]
marital_status	Married Binary	-0.3300***
		[0.0674]
num_dep	Number of Dependents	0.0630***
		[0.0194]
sex	Male=1	1.4722***
		[0.1006]
cd	CD binary	0.9581***
		[0.0624]
act	ACT Score	0.0011
		[0.0008]
lsat	LSAT Score	0.0019***
		[0.0007]
sat	SAT Score	-0.0012
		[0.0010]
rifle_class	Expert=1, 0 Otherwise	1.6040***
		[0.0604]
afqt	AFQT_SCORE	-0.0239***
		[0.0006]
dlab	DLAB Score	0.0307***
		[0.0006]
pft_class	1st Class=1, 0 Otherwise	-0.3392***
		[0.0782]
swim	Advanced =1, 0 Otherwise	0.0544
		[0.0587]
weight_control	1=ever assigned to weight control	-0.4312
		[0.2832]
xact	Missing ACT	-0.5738***
		[0.1458]
xlsat	Missing LSAT	1.1473***
		[0.2839]
_sat	Missing SAT	-1.7126***
		[0.4901]
sat_sat	sat * _sat	0.0052***
		[0.0010]
xdlab	Missing DLAB	0.7648***
		[0.2248]
xpft_class	Missing PFT Class	-0.2518
		[0.8375]
xswim	Missing Swim Class	0.4631
		[0.3171]
combat	Combat MOS Group	-3.7910***
		[0.0613]
non_combat	Non-Combat MOS Group	-5.0365***
		[0.0544]

VARIABLES	LABELS	Model 7
Constant	Constant	122.1023***
		[0.5587]
Observations		172,771
R-squared		0.093
Standard errors in brackets		
*** p<0.01, ** p<0.05, * p<0.1		

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX E. DESCRIPTIVE STATISTICS

Variable	Obs	Mean	Std. Dev.	Min	Max
fy	172771	2009.986	3.325674	2004	2015
marital_status	172771	0.731535	0.443162	0	1
num_dep	172771	1.827546	1.595326	0	13
mos	172771	1.070839	0.870985	0	2
sex	172771	0.943046	0.231755	0	1
cd	172771	0.678476	0.467063	0	1
act	172771	1.616562	29.04079	0	1520
lsat	172771	2.762275	40.44129	0	1560
sat	172771	1550.789	161.439	0	1600
rifle_class	172771	0.534737	0.498793	0	1
afqt	172771	23.7724	37.33802	0	99
dlab	172771	23.24714	43.14962	0	970
pft_class	172771	0.326583	0.468965	0	1
swim	172771	0.754982	0.430099	0	1
weight_control	172771	0.006303	0.079142	0	1
avg_hi	172771	-0.32453	0.408255	-3.66889	0.345
avg_this_rpt	172771	2.977357	1.939015	0	7
xact	172771	0.9715	0.166397	0	1
xlsat	172771	0.988783	0.105316	0	1
_sat	172771	0.995184	0.069228	0	1
sat_sat	172771	1549.572	170.2151	0	1600
xrifle	172771	0.281772	0.449864	0	1
xdlab	172771	0.044707	0.20666	0	1
xpft_class	172771	0.561495	0.496205	0	1
xswim	172771	0.005603	0.074642	0	1
gct	172771	124.9104	9.764524	42	160
buildup	172771	0.439293	0.496302	0	1
drawdown	172771	0.560708	0.496302	0	1
control_lawyers	172771	0.033235	0.17925	0	1
control_pilot	172771	0.31316	0.46378	0	1
combat	172771	0.236371	0.424854	0	1
non_combat	172771	0.417234	0.493104	0	1
combat_drawdown	172771	0.133581	0.340203	0	1
non_combat_drawdown	172771	0.239236	0.426618	0	1
combat_buildup	172771	0.102789	0.303684	0	1
non_combat_buildup	172771	0.177999	0.382513	0	1

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Bacolod, M. B. (2007). Do alternative opportunities matter? The role of female labor markets in the decline of teacher quality. *The Review of Economics and Statistics, 89*(4), 737–751.

Belasco, A. (2009). *Troop levels in the Afghan and Iraq wars, FY2001-FY2012: Cost and other potential issues* (CRS Report No. R40682). Retrieved from https://www.fas.org/sgp/crs/natsec/R40682.pdf.

Cancian, M. F., & Klein, M. W. (2015, July 20). Military officer quality in the all-volunteer force [Paper]. Retrieved from http://www.brookings.edu/research/papers/2015/07/20-nber-military-officer-quality-volunteer-force--klein

Chalabi, M. (2015, March 19). What percentage of Americans have served in the military? Retrieved from http://fivethirtyeight.com/datalab/what-percentage-of-americans-have-served-in-the-military/

Congressional Budget Office. (1999, November). *The drawdown of the military officer corps.* Washington, DC: Author. Retrieved from https://www.cbo.gov/sites/default/files/106th-congress-1999-2000/reports/drawdown.pdf.

deGrandpre, A. (2013, March 4). General: With cuts, Marine Corps will 'cut into bone.' *USA Today.* Retrieved from http://www.usatoday.com/story/news/nation/2013/03/03/marine-corps-amos-letter-sequester/1961117/

Department of Defense. (2014, October 30). *Commissioned Officer Promotion Reports (COPRs)* (DOD Instruction 1320.13). Retrieved from http://www.dtic.mil/whs/directives/corres/pdf/132013p.pdf.

Duflo, E. (n.d.). *Empirical Methods* [Lecture Notes]. Retrieved from http://web.mit.edu/14.771/www/emp_handout.pdf.

Garza, R. P. (2014). *United States Marine Corps career designation board: Significant factors in predicting selection* (Master's thesis). Retrieved from Calhoun http://calhoun.nps.edu/handle/10945/41381

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2011). *Impact evaluation in practice*. Retrieved from http://siteresources.worldbank.org/EXTHDOFFICE/Resources/5485726-1295455628620/Impact_Evaluation_in_Practice.pdf

Johnson, J. A. (2015). *Significant pre-accession factors predicting success or failure during a Marine Corps officer's initial service obligation.* Manuscript submitted for publication at Naval Postgraduate School.

Johnston, A. C. (2015). *The effects of personnel stability on Marine Corps Readiness: Are infantry battalions ready to respond to future conflicts?* Manuscript submitted for publication at Naval Postgraduate School.

Lamothe, D. (2010, July 13). Drastic cuts outlined in think-tank report [Web forum post]. Retrieved from http://www.ar15.com/forums/t_1_5/1065524__ARCHIVED_THREAD____Drastic_Military_cuts_outlined_in_think_tank_report.html&page=1

Manpower Management Records and Performance Branch (MMRP-30). (n.d.). Manpower & Reserve Affairs. Retrieved December 17, 2015, from https://www.manpower.usmc.mil/portal/page/portal/M_RA_HOME/MM/G_SB/b_MMSB_30_PERFORMANCE_EVALUATION

Marine Corps University. (n.d.). *Regional, culture, and language familiarization* [Interactive map]. Retrieved December 28, 2015, from http://www.mcu.usmc.mil/caocl/SitePages/rclf.aspx

Rostker, B. (2013). *Right-sizing the force: Lessons learned for the current drawdown of American military personnel*. Retrieved from https://www.cnas.org/sites/default/files/publications-pdf/CNAS_RightSizingTheForce_Rostker_0.pdf

Shinkman, P. D. (2013, May 23). Obama: ‘Global war on terror’ is over. *U.S. News & World Report.* Retrieved from http://www.usnews.com/news/articles/2013/05/23/obama-global-war-on-terror-is-over

Sisk, R. (2015, February 2). Marine Corps proposes $25B budget keeping end strength at 184k. Retrieved from http://www.military.com/daily-news/2015/02/02/marine-corps-proposes-25b-budget-keeping-end-strength-at-184k.html.

Sneden, P. (2009, December 15). 202K. Retrieved from http://marinesmagazine.dodlive.mil/2009/12/15/202k/.

Timelines of the Iraq and Afghanistan wars. (2013). *NewsOK*. Retrieved from http://newsok.com/article/3770937

Total Forces Data Warehouse (TFDW). (n.d.). Total Force Data Warehouse. Retrieved December 17, 2015, from https://tfdw-web.mceits.usmc.mil/index.aspx

United States Marine Corps. (n.d.). The Basic School training command. Retrieved December 28, 2015, from http://www.trngcmd.marines.mil/Units/Northeast/TheBasicSchool.aspx

United States Marine Corps. (2008a). *Career designation, retention, and return to active duty, redesignation of restricted officers to unrestricted status, and interservice transfer of officers into the Marine Corps* (MCO 1001.45J). Retrieved from http://www.marines.mil/Portals/59/Publications/MCO%201001.45J.pdf.

United States Marines Corps. (2008b). *Marine Corps body composition and military appearance program* (MCO 6110.3). Retrieved from http://hprc-online.org/physical-fitness/files/usmc-mco-6110.3-percent-body-fat.

United States Marine Corps. (2008c). *Marine Corps end strength increase* (ALMAR 008/07). Retrieved from http://www.marines.mil/News/Messages/Messages Display/tabid/13286/Article/109650/marine-corps-end-strength-increase.aspx.

United States Marine Corps. (2010a). *Fiscal year 2010 officer retention board career designation, return to competitive career designation and change in meritorious career designation policy* (MARADMIN 021/09) Retrieved from http://www.marines.mil/News/Messages/MessagesDisplay/tabid/13286/Article/112192/fiscal-year-2010-officer-retention-board-career-designation-return-to-competiti.aspx

United States Marine Corps. (2010b). *Marine Corps water survival training program* (MCWSTP) (MCO 1500.52D). Retrieved from http://www.marines.mil/Portals/59/Publications/MCO%201500.52D.pdf

United States Marine Corps. (2015). *Performance evaluation system (PES)* (MCO 1610.7). Retrieved from http://www.marines.mil/Portals/59/Publications/MCO%201610.7.pdf

United States Marine Corps History Division. (2014). Marine Corps end strengths. Retrieved from http://www.mcu.usmc.mil/historydivision/pages/frequently_requested/EndStrength.aspx

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California